IN THE UNITED STATES PATENTS AND TRADE MARK OFFICE

Applicant:	Commonwealth Scientific and Industrial Research Organisation
Serial Number:	09/776,910
Filed:	6 February 2001
Titled:	Malathion Carboxylesterase

DECLARATION UNDER 37 C.F.R. 1.132

I, **Robyn Russell** of Canberra, Australia declare that:

1. I am one of the inventors of the subject matter of U.S. Patent Application Serial No. 09/776,910 (the present application) which is a divisional of U.S. Patent Application Serial No. 09/068,960 (now U.S. 6,235,515) originally filed as PCT/AU96/00746 filed on 22 November 1996 which claims priority from Australian Provisional Application No. PN 6751 filed on 23 November 1995.

2. My qualifications and technical experience are set out in my Curriculum Vitae, a copy of which is attached as Annexure A.

3. The present application relates to the isolation and characterisation of esterases with the ability to hydrolyse organophosphates such as malathion.

4. It is my understanding that the following claim is pending in the present application:

   9. *A recombinant enzyme capable of hydrolyzing at least one organophosphate selected from the group consisting of carboxylester organophosphates and dimethyl-oxon organophosphates, wherein the recombinant enzyme comprises an amino acid sequence which is at least about 75% identical SEQ ID NO:8, wherein the recombinant enzyme comprises amino acid residue conserved between the sequences provided in Figure 4 with the exception that the recombinant enzyme comprises an amino acid selected from the group consisting of Leu, Ser, Ala, Ile, Val, Thr, Cys, Met and Gly at position 251.*

5. As outlined in the present application, myself and co-inventors endeavoured to identify an enzyme which confers organophosphate resistance upon the Dipteran *Lucilia cuprina*. Following a number of failed attempts, myself and co-inventors isolated and characterised a molecule which was found to be a mutant of an already known esterase referred to in the art as "E3".

6. Since the present application was filed myself and co-workers have made a further protein encompassed by pending claim 9. In this protein, the naturally occurring serine at position 250 has been replaced with a proline. This mutant possesses organophosphate hydrolysing activity (see Table below).

Enzyme	**Titre** (pmol μl^{-1})	**dECP** k_{cat} (min^{-1})	**dMUP** k_{cat} (min^{-1})
E3W251L	6.9 ± 2.0	0.0092 ± 0.0004	0.0610 ± 0.0060
E3P250S/W251L	6.7 ± 0.80	0.0036 ± 0.0003	0.0180 ± 0.0000

7. My laboratory has also made a mutant of the protein from *Drosophila melanogaster* which is orthologous to the *Lucilia cuprina* and *Musca domestica* proteins disclosed in the present application. The mutated version of the *D. melanogaster* protein is **63% identical** to the claimed *L. cuprina* molecule, and 76% identical to the *M. domestica* protein, and maintains the biological activity defined in pending claim 9. Alignments of these proteins are provided in Annexure B. A comparison of the biological activity is provided in the following Table.

Enzyme	Titre (pmol μl^{-1})	dECP k_{cat} (min^{-1})	dMUP k_{cat} (min^{-1})
E3W251L	6.9 ± 2.0	0.0092 ± 0.0004	0.0610 ± 0.0060
EST23W251L*	0.9	0.0060 ± 0.0004	0.0210 ± 0.0003

* EST23 is the designation of the *D. melanogaster* ortholog of *L. cuprina* E3

8. My laboratory has made many other mutants of the protein provided as SEQ ID NO: 8 in the present application. I do not recall producing any mutants which are encompassed by pending claim 9 which lacked organophosphate hydrolysing activity.

9. As outlined herein, the proteins described in the present application are esterases. Esterases form a large protein family with multiple esterases with varying activity being found in many different organisms such as animals (including insects and mammals), plants, fungi and bacteria.

10. At the earliest priority date, namely 23 November 1995, it is my opinion that the structure/function relationship of various esterases was well understood. In particular, important functional domains had been identified, and the skilled person was well aware of important conserved residues.

11. With regard to Dipterans, our understanding of different esterases in this family of organisms was also well advanced at 23 November 1995. For instance, a cluster of esterases had already been identified and characterised from another Dipteran, namely *Drosophila melanogaster*. A copy of publications describing this large family of esterases, and structural/functional characteristics thereof) is attached as Annexure C (namely, Oakeshott, J.G., van Papenrecht, E.A., Boyce, T.M., Healy, M.J. and Russell, R.J. (1993) Evolutionary genetics of *Drosophila* esterases. Genetica 90: 239-268, and Karotam, J., Delves, A.C. and Oakeshott, JG. (1993) Conservation and change in structural and 5' flanking sequences of esterase 6 in sibling *Drosophila* sequences. Genetica 88: 11-28).

Dated this ....10 th day of Feb..... 2005

Robyn J. Russell

Robyn Russell

ALIGNMENT OF *LUCILIA CUPRINA*, *MUSCA DOMESTICA* AND *DROSOPHILA MELANOGASTER* PROTEINS

```
Lc      MNFNVSLMEKLKWKIKCIENKFLNYRLTTNETVVAETEYGKVKGVKRLTVYDDSYYSFEG   60
Md      MNFKVSQMERLSWKLKCMVNKYTNYRLSTNETQIIDTEYGQIKGVKRMTVYDDSYYSFES   60
Dm      MNKNLGFVERLRWRLKTIEHKVQQYRQSTNETVVADTEYGQVRGIKRLSLYDVPYFSFEG   60
        ** ::. :*:* *::* : :*  :** :**** : :****:::*:**:::** .*:***.

Lc      IPYAQPPVGELRFKAPQRPTPWDGVRDCCNHKDKSVQVDFITGKVCGSEDCLYLSVYTNN   120
Md      IPYAKPPVGELRFKAPQRPVPWEGVRDCCGPANRSVQTDFISGKPTGSEDCLYLNVYTND   120
Dm      IPYAQPPVGELRFKAPQRPIPWERVRDCSQPKDKAVQVQFVFDKVEGSEDCLYLNVYTNN   120
        ****:*************** **: ****.   :::**.:*: .*  ********.****:

Lc      LNPETKRPVLVYIHGGGFIIGENHRDMYGPDYFIKKDVVLINIQYRLGALGFLSLNSEDL   180
Md      LNPDKKRPVMVFIHGGDFIFGEANRNWFGPDYFMKKPVVLVTVQYRLGVLGFLSLKSENL   180
Dm      VKPDKARPVMVWIHGGGFIIGEANREWYGPDYFMKEDVVLVTIQYRLGALGFMSLKSPEL   180
        ::*:. ***:*:****.**:** :*: :*****:*: ***:.:*****.***:**:* :*

Lc      NVPGNAGLKDQVMALRWIKNNCANFGGNPDNITVFGESAGAASTHYMMLTEQTRGLFHRG   240
Md      NVPGNAGLKDQVMALRWVKSNIANFGGDVDNITVFGESAGGASTHYMMITEQTRGLFHRG   240
Dm      NVPGNAGLKDQVLALKWIKNNCASFGGDPNCITVFGESAGGASTHYMMLTDQTQGLFHRG   240
        ************:**:*:*.* *.***: : *********.*******:*:**:******

Lc      ILMSGNAICPLANT-QCQHRAFTLAKLAGYKGEDNDKDVLEFLMKAKPQDLIKLEEKVLT   299
Md      IMMSGNSMCSLAST-ECQSRALTMAKRVGYKGEDNEKDILEFLMKANPYDLIKEEPQVLT   299
Dm      ILQSGSAICPLAYNGDITHNPYRIAKLVGYKGEDNDKDVLEFLQNVKAKDLIRVEENVLT   300
        *: **.::*.** . :   .. :** .*******:**:**** :.:. ***: * :***

Lc      LEERTNKVMFPFGPTVEPYQTADCVLPKHPREMVKTAWGNSIPTMMGNTSYEGLFFTSIL   359
Md      PEEMQNKVMFPFGPTVEPYQTADCVVPKPIREMVKSAWGNSIPTLIGNTSYEGLLFKSIA   359
Dm      LEERMNKIMFAFGPSLEPFSTPECVISKPPKEMMKTAWSNSIPMFIGNTSYEGLLWVPEV   360
         **  **:**.***::**:.*.:**:.*  :**:*:**.**** ::********:: .

Lc      KQMPMLVKELETCVNFVPSELADAERTAPETLEMGAKIKKAHVTGETPTADNFMDLCSHI   419
Md      KQYPEVVKELESCVNYVPWELADSERSAPETLERAAIVKKAHVDGETPTLDNFMELCSYF   419
Dm      KLMPQVLQQLDAGTPFIPKELLATEPSKEKLDSWSAQIRDVHRTGSESTPDNYMDLCSIY   420
        *  * ::::*:: . ::* **  :* :  :  . .* ::..*  *. .* **:*:***

Lc      YFWFPMHRLLQLRFNHTSGTPVYLYRFDFDSEDLINPYRIMRSGRGVKGVSHADELTYFF   479
Md      YFLFPMHRFLQLRFNHTAGTPIYLYRFDFDSEEIINPYRIMRFGRGVKGVSHADELTYLF   479
Dm      YFVFPALRVVHSRHAYAAGAPVYFYRYDFDSEELIFPYRIMRLGRGVKGVSHADDLSYQF   480
        ** **  *.:: *. :::*:*:*:**:*****:* ****** **********:*:* *

Lc      WNQLAKRMPKESREYKTIERMTGIWIQFATTGNPYSNEIEGMENVSWDPIKKSDEVYKCL   539
Md      WNILSKRLPKESREYKTIERMVGIWTEFATTGKPYSNDIAGMENLTWDPIKKSDDVYKCL   539
Dm      SSLLARRLPKESREYRNIERTVGIWTQFAATGNPYSEKINGMDTLTIDPVRKSDEVIKCL   540
         . *::*:*******:.*** .*** :**:**:***:.* **:.:: **::***:* ***

Lc      NISDELKMIDVPEMDKIKQWESMFEKHRDLF-                               570
Md      NIGDELKVMDLPEMDKIKQWASIFDKKKELF-                               570
Dm      NISDDLKFIDLPEWPKLKVWESLYDDNKDLLF                               572
        **.*:**.:*:**  *:* * *:::.:::*:
```

ALIGNMENT OF *LUCILIA CUPRINA* AND *DROSOPHILA MELANOGASTER* PROTEINS

```
Lc      MNFNVSLMEKLKWKIKCIENKFLNYRLTTNETVVAETEYGKVKGVKRLTVYDDSYYSFEG   60
Dm      MNKNLGFVERLRWRLKTIEHKVQQYRQSTNETVVADTEYGQVRGIKRLSLYDVPYFSFEG   60
        ** *:.:::*:*:*::* **:*. :** :*******:****:*:*:***::** .*:****

Lc      IPYAQPPVGELRFKAPQRPTPWDGVRDCCNHKDKSVQVDFITGKVCGSEDCLYLSVYTNN  120
Dm      IPYAQPPVGELRFKAPQRPIPWERVRDCSQPKDKAVQVQFVFDKVEGSEDCLYLNVYTNN  120
        ******************** **: ****.: ***:***:*: .** ********.*****

Lc      LNPETKRPVLVYIHGGGFIIGENHRDMYGPDYFIKKDVVLINIQYRLGALGFLSLNSEDL  180
Dm      VKPDKARPVMVWIHGGGFIIGEANREWYGPDYFMKEDVVLVTIQYRLGALGFMSLKSPEL  180
        ::*:. ***:*:********** :*: ******:*:****:.**********:**:* :*

Lc      NVPGNAGLKDQVMALRWIKNNCANFGGNPDNITVFGESAGAASTHYMMLTEQTRGLFHRG  240
Dm      NVPGNAGLKDQVLALKWIKNNCASFGGDPNCITVFGESAGGASTHYMMLTDQTQGLFHRG  240
        ************:**:*******.***:*: *********.*********:**:******

Lc      ILMSGNAICPLANT-QCQHRAFTLAKLAGYKGEDNDKDVLEFLMKAKPQDLIKLEEKVLT  299
Dm      ILQSGSAICPLAYNGDITHNPYRIAKLVGYKGEDNDKDVLEFLQNVKAKDLIRVEENVLT  300
        ** **.****** . :  *..: :***.*************** :.*.:***::**:***

Lc      LEERTNKVMFPFGPTVEPYQTADCVLPKHPREMVKTAWGNSIPTMMGNTSYEGLFFTSIL  359
Dm      LEERMNKIMFAFGPSLEPFSTPECVISKPPKEMMKTAWSNSIPMFIGNTSYEGLLWVPEV  360
        **** **:**.***::**:.*.:**:.* *:**:****.**** ::********::.. :

Lc      KQMPMLVKELETCVNFVPSELADAERTAPETLEMGAKIKKAHVTGETPTADNFMDLCSHI  419
Dm      KLMPQVLQQLDAGTPFIPKELLATEPSKEKLDSWSAQIRDVHRTGSESTPDNYMDLCSIY  420
        * ** ::::*:: . *:*.**  :* :  :  . .*:*:..* **. .*.**:*****

Lc      YFWFPMHRLLQLRFNHTSGTPVYLYRFDFDSEDLINPYRIMRSGRGVKGVSHADELTYFF  479
Dm      YFVFPALRVVHSRHAYAAGAPVYFYRYDFDSEELIFPYRIMRLGRGVKGVSHADDLSYQF  480
        ** **  *::: *. :::*:***:**:*****:** ****** **********:*:* *

Lc      WNQLAKRMPKESREYKTIERMTGIWIQFATTGNPYSNEIEGMENVSWDPIKKSDEVYKCL  539
Dm      SSLLARRLPKESREYRNIERTVGIWTQFAATGNPYSEKINGMDTLTIDPVRKSDEVIKCL  540
         . **:*:*******:.*** .*** ***:******::*:**:.:: **::***** ***

Lc      NISDELKMIDVPEMDKIKQWESMFEKHRDLF-                              570
Dm      NISDDLKFIDLPEWPKLKVWESLYDDNKDLLF                              572
        ****:**:**:**  *:* ***:::.::**:
```

CURRICULUM VITAE

Full Name: Robyn Joyce Russell

Date of Birth: June 23, 1952

Citizenship: Australian

Marital Status: Married, to J.G. Oakeshott, two children

Work Address: CSIRO Entomology
GPO Box 1700
Canberra, ACT 2601
Australia
Telephone: (02) 6246 4160
Fax: (02) 6246 4173
Email: Robyn.Russell@ento.csiro.au

Home Address: 28 Sullivan Crescent
Wanniassa
Canberra, ACT 2903
(02) 6231 4196

Qualifications:

1975-79	Melbourne University, Parkville, Victoria, Australia, PhD in Viral Immunology
1970-74	Monash University, Clayton, Victoria, Australia, BSc (Hons IIA) in Biochemistry

Awards and Honours:

1974	Commonwealth Postgraduate Research Award
1970	Commonwealth Tertiary Scholarship
1967	Commonwealth Secondary Scholarship
1965	Victorian State Government Scholarship

Professional Experience:

1996-present	Principal Research Scientist CSIRO Entomology GPO Box 1700, Canberra, ACT 2601, Australia

1987-96	Senior Research Scientist CSIRO Entomology GPO Box 1700, Canberra, ACT 2601, Australia
1985-87	Experimental Scientist CSIRO Wildlife and Rangelands Research PO Box 84, Lyneham, ACT 2602, Australia
1982-85	Research Fellow Department of Population Biology Research School of Biological Sciences The Australian National University, Canberra, ACT, Australia
1979-81	Postdoctoral Fellow Department of Microbiology The Pennsylvania State University College of Medicine Hershey, Pennsylvania, USA

Administrative Experience:

2001-present	Chair, Institutional Biosafety Committee, CSIRO Entomology
1989-2002	Member, Institutional Biosafety Committee, CSIRO Entomology
1993-97	Member, Institutional Biosafety Committee CSIRO Wildlife and Ecology
1988-91	Member, Committee for the Reclassification of Technical Staff, CSIRO Entomology
1985-87	Member, Institutional Biosafety Committee, CSIRO Wildlife and Rangelands Research

Undergraduate Teaching Experience:

1998	Honours Supervisor, Robert McCuaig Biochemistry and Molecular Biology, ANU
1997	Honours Supervisor, Karen Bell (ClassI) Biological Science and Medicine, Flinders University
1996	Honours Supervisor, Jeremy Brownlie (Class I) Botany and Zoology, ANU
1993-94	Honours Supervisor, Lyndall Briggs (Class IIA)

	Botany and Zoology, ANU
1992-93	Honours Supervisor, Leon Court (Class IIA) Botany and Zoology, ANU
1991-92	Honours Supervisor, Paris Kostakos (Class IIA) Botany and Zoology, ANU
1990	Honours Supervisor, M. Spackman (Class 1), Genetics, ANU
1988-present	Guest Lecturer, Molecular Genetics (The Australian National University)
1979-81	Demonstrator (6 hours/week), Medical Virology (The Pennsylvania State University College of Medicine)

Doctoral Student Supervision:

K. Weir	(Enrolled 2001). Bioremediation of pesticide residues in irrigation drainage waters. Co-supervisor.
E. Crone	(Enrolled 1999). Cloning and characterisation of juvenile hormone esterase gene in *Drosophila melanogaster*. Co-supervisor.
R. Heidari	(Enrolled 1998). Bioremediation of pesticide residues in irrigation drainage waters. Co-supervisor.
C. Claudianous	(Enrolled 1994). Molecular analysis of an α–esterase gene cluster in *Musca domestica*. Co-supervisor.
G.C. Robin	(Enrolled 1992). Molecular analysis of an α–esterase gene cluster on chromosome 3R of *Drosophila melanogaster*. Co-supervisor.
R.D. Newcomb	(Enrolled 1991). Molecular cloning of esterase genes involved in organophosphate insecticide resistance in *Lucilia cuprina*. Co-supervisor.
K.A. Smyth	(Enrolled 1990). Molecular basis of malathion insecticide resistance in *Lucilia cuprina*. Co-supervisor.
A. Parker	(Enrolled 1989). Molecular analysis of organophosphate insecticide resistance in *Lucilia cuprina*. Co-supervisor.
J. Karotam	(Awarded 1993). Conservation and change in *Esterase 6* nucleotide sequences of *Drosophila*. Co-supervisor.

P. Christian	(Awarded 1988). Studies on *Drosophila* C and A viruses in Australian populations of *Drosophila melanogaster*. Co-supervisor.
M. Healy	(Awarded 1985). Molecular and genetic studies of the uncoordinated gene of *Drosophila melanogaster*. Co-supervisor.

PUBLICATIONS — R.J. RUSSELL (1978-2004)

Refereed Journal Articles

1. Russell, R.J. and Jackson, D.C. (1978). Direct solid-phase radioimmunoassay for measuring antigenic differences between the haemagglutinins of influenza viruses. Journal of Immunological Methods, *22*: 201-209.

2. Jackson, D.C., Russell, R.J., Ward, C.W. and Dopheide, T.A. (1978). Antigenic determinants of influenza virus haemagglutinin. I. Cyanogen bromide peptides derived from A/Memphis/72 haemagglutinin possess antigenic activity. Virology, *89*: 199-205.

3. Jackson, D.C., Dopheide, T.A., Russell, R.J., White, D.O. and Ward, C.W. (1979). Antigenic determinants of influenza virus haemagluttinin. II. Antigenic activity of the isolated N-terminal cyanogen bromide peptide of A/Memphis/72 haemagglutinin heavy chain. Virology, *93*: 458-465.

4. Russell, R.J., Burns, W.H., White, D.O., Anders, E.M., Ward, C.W. and Jackson D.C. (1979). Antigenic determinants of influenza virus haemagluttinin. III. Competitive binding of antibodies directed against "common" and "strain-specific" antigenic determinants of A/Memphis/72 haemagglutinin. Journal of Immunology, *123*: 825-832.

5. Miller, R.H., Russell, R.J. and Hyman, R.W. (1982). Physical map of the short foldback sequences of herpes simplex virus type I DNA. Virology, *117*: 70-80.

6. Russell, R.J., Kudler, L., Miller, R.H. and Hyman, R.W. (1982). Stability of the cloned 'joint region' of herpes simplex virus DNA. Intervirology, *18*: 98-104.

7. Kudler, L., Jones, T.R., Russell, R.J. and Hyman, R.W. (1983). Hetroduplex analysis of cloned fragments of herpes simplex virus DNAs. Virology, *124*: 86-93.

8. Miklos, G.L.B., Healy, M.J., Pain, P., Howells, A.J. and Russell, R.J. (1984). Molecular and genetic studies on the euchromatin-heterochromatin transition region of the X chromosome of *Drosophila melanogaster* I. A cloned entry point near to the *uncoordinated* locus. Chromosoma, *89*: 218-227.

9. Oakeshott, J.G., Collet, C., Phillis, R., Nielsen, K.M., Russell, R.J., Chambers, G.K., Ross, V. and Richmond, R.C. (1987). Molecular cloning and characterisation of Esterase 6, a serine hydrolase from *Drosophila*. Proceedings of the National Academy of Sciences U.S.A. *84*: 3359-3363.

10. Healy, M.J. Russell, R.J. and Miklos, G.L.G. (1988). Molecular and genetic studies on the euchromatin-heterochromatin transition region of the X chromosome of *Drosophila melanogaster* II. Isolation of the *uncoordinated* gene. Molecular and General Genetics, *213*: 63-71.

11. Russell, R.J. and Robbins, S.J. (1989). Cloning and molecular characterisation of the myxoma virus genome. Virology, *170*: 147-159.

12. Collet, C., Nielsen, K.M., Russell, R.J., Karl, M., Collet, S., Oakeshott, J.G. and Richmond, R.C. (1990). Molecular analysis of duplicated esterase genes in *Drosophila melanogaster*. Molecular Biology and Evolution, *7*: 9-28.

13. Healy, M.J., and Russell, R.J. (1990). Transcriptional analysis of the *uncoordinated* region of *Drosophila melanogaster*. Genome, *33*: 829-836.

14 Parker, A.G., Russell, R.J., Delves, A.C. and Oakeshott, J.G. (1991). Biochemistry and physiology of esterases in organophosphate susceptible and resistant strains of the Australian sheep blowfly, *Lucilia cuprina*. Pesticide Biochemistry and Physiology, *41*: 305-318.

15. Russell, R.J., Healy, M.J. and Oakeshott, J.G. (1992). Molecular analysis of the *lethal (1)B214* region at the base of the X chromosome of *Drosophila melanogaster*. Chromosoma, *101*: 456-466.

16. Oakeshott, J.G., van Papenrecht, E.A., Boyce, T.M., Healy, M.J. and Russell, R.J. (1993). Evolutionary genetics of *Drosophila* esterases. Genetica *90*: 239-268.

17. Spackman, M.E., Oakeshott, J.G., Smyth, K-A., Medveczky, K.M. and Russell, R.J. (1994). A cluster of esterase genes on chromosome 3R of *Drosophila melanogaster* includes homologues of esterase genes conferring insecticide resistance in *Lucilia cuprina*. Biochemical Genetics *32*: 39-62.

18. Board, P., Russell, R.J., Marano, R.J. and Oakeshott, J.G. (1994). Purification, molecular cloning and heterologous expression of a glutathione S-transferase from the Australian sheep blowfly *Lucilia cuprina*.. Biochemical Journal *299*: 425-430.

19. Whyard, S., Russell, R.J. and Walker, V.K. (1994). Insecticide resistance and malathion carboxylesterase in the sheep blowfly, *Lucilia cuprina*.. Biochemical Genetics *32*: 9-24.

20. Smyth, K-A., Russell, R.J. and Oakeshott, J.G. (1994) A cluster of at least three esterase genes in *Lucilia cuprina* includes malathion carboxylesterase and two other esterases implicated in resistance to organophosphates. Biochemical Genetics *32*: 437-453.

21. Oakeshott, J.G., Boyce, T.M., Russell, R.J. and Healy, M.J. (1995). Molecular insights into the evolution of an enzyme; esterase6 in *Drosophila*. Trends in Ecology and Evolution *10*: 103-110.

22. Oakeshott, J.G., Boyce, T.M., Russell, R.J. and Healy, M.J. (1995). Reply to: Dover, G. Molecular coevolution and the esterase enigma. Trends in Ecology and Evolution *10*: 287.

23. Russell, R.J., Robin, C., Kostakos, P., Newcomb, R.D., Boyce, T.M., Medveczky, K.M. and Oakeshott, J.G. (1996). Molecular cloning of an *a*-esterase gene cluster on chromosome 3R of *Drosophila melanogaster.* Insect Biochemistry and Molecular Biology *26* : 235-247.

24. Smyth, K-A., Walker, V.K., Russell, R.J. and Oakeshott, J.G. (1996). Biochemical and physiological differences in the malathion carboxylesterase activities of malathion susceptible and resistant lines of the sheep blowfly, *Lucilia cuprina.* Pesticide Biochemistry and Physiology *54* : 48-55.

25. Parker, A.G., Campbell, P.M., Spackman, M.E., Russell, R.J. and Oakeshott, J.G. (1996). Comparison of an esterase associated with organophosphate resistance in *Lucilia cuprina* with an orthologue not associated with resistance in *Drosophila melamogaster.* Pesticide Biochemistry and Physiology *55* : 85-99.

26. Newcomb, R.D., East, P.D., Russell, R.J., and Oakeshott, J.G. (1996). Isolation of α-cluster esterase genes associated with organophosphate resistance in *Lucilia cuprina* . Insect Molecular Biology *5* : 211-216.

27. Robin, C., Russell, R.J., Medveczky, K.M. and Oakeshott, J.G. (1996). Duplication and divergence of the genes of the α–esterase cluster of *Drosophila melanogaster.* Journal of Molecular Evolution *43* : 241-252.

28. Newcomb, R.D., Campbell, P.M., Russell, R.J. and Oakeshott, J.G. (1997). cDNA cloning, baculovirus expression and kinetic properties of the esterase, E3, involved in organophosphate resistance in *Lucilia cuprina.* Insect Biochemistry and Molecular Biology *27* : 15-25.

29. Campbell, P.M., Trott, J.F., Claudianos, C., Smyth, K-A., Russell, R.J. and Oakeshott, J.G. (1997). Biochemistry of esterases associated with organophosphate resistance in *Lucilia cuprina* with comparisons to putative orthologues in other Diptera. Biochemical Genetics *35* : 17-39.

30. Newcomb, R.D., Campbell, P.M., Ollis, D.L., Cheah, E., Russell, R.J. and Oakeshott, J.G. (1997) A single amino acid substitution converts a carboxylesterase to an organophosphorus hydrolase and confers insecticide resistance on a blowfly. PNAS *94* : 7464-7468.

31. Dumancic, M.M., Oakeshott, J.G., Russell, R.J. and Healy, M.J. (1997) Characterisation ot the *EstP* protein of *Drosophila melanogaster* and its conservation in Drosophilids. Biochemical Genetics *35:* 251-271.

32. Campbell, P.M., Newcomb, R.D., Russell, R.J. and Oakeshott, J.G. (1998) Two different amino acid substitutions in the ali-esterase, E3, confer alternative types of organophosphorus insecticide resistance in the sheep blowfly, *Lucilia cuprina.* Insect Biochemistry and Molecular Biology *28:* 139-150.

33. Campbell, P.M., Yen, J.L., Masoumi, A., Russell, R.J., Batterham, P., McKenzie, J.A. and Oakeshott, J.G. (1998) Cross-resistance patterns among Australian sheep blowfly, *Lucilia cuprina* (Diptera: Calliphoridae), resistant to organophosphorus insecticides. Journal

of Economic Entomology *91:* 367-375.

34. Smyth, K-A., Russell, R.J. and Oakeshott, J.G. (1998) Propetamphos resistance in the Australian sheep blowfly, *Lucilia cuprina* (Diptera: Calliphoridae). Australian Journal of Entomology *37:* 57-59.

35. Claudianos, C., Russell, R.J. and Oakeshott, J.G. (1999) The same amino acid substitution in orthologous esterases confers organophosphate resistance on the housefly and a blowfly. Insect Biochemistry and Molecular Biology. *29:* 675-686.

36. Oakeshott, J.G., Claudianos, C., Russell, R.J. and Robin, G.C. (1999) Carboxyl / Cholinesterases: A case study of the evolution of a successful multigene family. *BioEssays 21*:1031-1042

37. Smyth, K-A., Boyce, T.M., Russell, R.J. and Oakeshott, J.G. (2000) MCE activities and malathion resistances in field populations of the Australian sheep blowfly (*Lucilia cuprina*). *Heredity. 84:* 63-72.

38. Sutherland, T., Horne, I., Lacey, M.J., Harcourt, R.L., Russell, R.J. and Oakeshott, J.G. (2000) Enrichment of an endosulfan-degrading mixed bacterial culture. *Applied and Environmental Microbiology. 66:* 2822-2828.

39. Robin, G.C., Russell, R.J., Cutler, D.J. and Oakeshott, J.G. (2000) The evolution of an α-esterase pseudogene inactivated in the *Drosophila melanogaster* lineage. *Molecular Biology and Evolution. 17:* 563-575.

40. Robin, G. C., Claudianos, C, Russell, R.J. and Oakeshott, J.G. (2000) Reconstructing the diversification of α-esterases; comparing the gene clusters of *Drosophila buzzatii* and *D. melanogaster. Journal of Molecular Evolution 51:* 149-160.

41. Campbell, P.M., Harcourt, R.L., Crone, E.J., Claudianos, C., Hammock, B.D., Russell, R.J. and Oakeshott, J.G. (2001) Identification of a juvenile hormone esterase gene by matching its peptide mass fingerprint with a sequence from the *Drosophila* genome project. *Insect Biochemistry and Molecular Biology 31:* 513-520.

42. Horne, I., R.L. Harcourt, T.D. Sutherland, R.J. Russell & J.G. Oakeshott. (2002) Isolation of a *Pseudomonas monteilli* strain with a novel phosphotriesterase. *FEMS Microbiol Lett. 206:* 51-55.

43. Harcourt, R.L., I. Horne, T.D. Sutherland, B.D. Hammock, R.J. Russell & J.G. Oakeshott. (2002) Development of a simple and sensitive fluorimetric method for isolation of coumaphos-hydrolysing bacteria. *Lett. Appl. Microbiol. 34*: 263-268.

44. Sutherland, TD, KM Weir, MJ Lacey, I Horne, RJ Russell, and JG Oakeshott. (2002) Enrichment of a bacterial culture capable of degrading endosulfate, the toxic metabolite of endosulfan. *J. Appl. Microbiol. 92:* 541-548.

45. Sutherland, TD, I Horne, RL Harcourt, RJ Russell, and JG Oakeshott. (2002) Isolation and characterisation of a *Mycobacterium* strain that metabolises the insecticide endosulfan. *J. Appl. Microbiol.* 93:380-389.

46. Horne, I., T.D. Sutherland, R.L. Harcourt, R.J. Russell & J.G. Oakeshott. 2002. Identification of an *opd* (organophosphate degradation) gene in an *Agrobacterium* isolate. *Appl. Environ. Microbiol.* 68: 3371-3376.

47. Horne, I., T.D. Sutherland, J.G. Oakeshott & R.J. Russell. (2002) Cloning and expression of the phosphotriesterase gene *hocA* from *Pseudomonas monteilli* C11. Microbiology; *148*: 2687-2695.

48. Claudianos, C., Brownlie, J., Russell, R., Oakeshott, J.G., and Whyard, S. (2002) *maT* - a clade of transposons intermediate between *mariner* and *Tc1*. *Molec. Biol. Evol.* 19: 2101-2109.

49. Sutherland, T.D., Horne, I., Russell, R.J and Oakeshott, J.G. (2002) Gene cloning and molecular characterization of a two enzyme system catalyzing the oxidative detoxification of beta-endosulfan. *Appl. Environ. Microbiol.* 68: 6237-6245.

50. Oakeshott, J.G., Horne, I., Sutherland, T.D., Russell, R.J. (2003) The genomics of insecticide resistance. *Genome Biology* 4: 202.1-202.4.

51. Yang, H., Carr, P.D., McLoughlin, S. Yu, Liu, J.W., Horne, I., Qiu, X. Jeffries, C.M.J., Russell, R.J., Oakeshott J.G. and Ollis D.L. (2003) Evolution of an organophosphate degrading enzyme; a comparison of natural and directed evolution. *Protein Engineering* 16: 135-145.

52. Devonshire, A.L., Heidari, R., Bell, K.L., Campbell, P.M., Campbell, B.E., Odgers, W.A., Oakeshott, J.G. and Russell, R.J. (2003) Kinetic efficiency of mutant carboxylesterases implicated in organophosphate insecticide resistance. *Pestic. Biochem. Physiol.* 76: 1-13.

53. Horne, I., Qiu, X., Russell R.J. and Oakeshott, J.G. (2003) The phosphotriesterase gene *opdA* in *Agrobacterium radiobacter* P230 is transposable. FEMS *Microbiology Letters* 222: 1-8.

54. Campbell, PM, de Q. Robin, GC, Court, LN, Dorrian, SJ, Russell, RJ and Oakeshott, JG. (2003) Developmental expression and gene/enzyme identifications in the alpha esterase gene cluster of *Drosophila melanogaster*. *Insect Molecular Biology* 12: 459-471.

55. Heidari, R, Devonshire, AL., Campbell, BE., Bell, KL., Dorrian, SJ., Oakeshott, JG. and Russell, RJ. (2004) Hydrolysis of organophosphorus insecticides by in vitro modified carboxylesterase E3 from *Lucilia cuprina*. *Insect Biochemistry and Molecular Biology.* 34: 353-363.

56. Sutherland, TD, Horne, I, Weir, KW, Russell, RJ and Oakeshott, JG. (2004) Evaluation of the toxicity and residues associated with the isomers of the insecticide endosulfan. *Reviews in Environmental Contamination and Toxicology.* 182: 99-113.

57. Russell, R.J., Claudianos, C., Campbell, P.M., Horne, I., Sutherland, T.D. and Oakeshott, J.G. Two major classes of target site insensitivity mutations confer resistance to organophosphate and carbamate insecticides. *Pesticide Biochemistry and Physiology.* 79: 84-93.

58. Sutherland, T.D., Horne, I., Weir, K.M., Coppin, C.W., Williams, M.R., Selleck, M., Russell, R.J. and Oakeshott, J.G. (2004) Enzymatic bioremediation: from enzyme discovery to applications. *Clinical and Experimental Pharmacology and Physiology.* 31: 817-821.

59. Newcomb, RD., Gleeson, DM., Yong, CX., Russell, RJ. and Oakeshott, JG. (2005) Multiple mutations and gene duplications conferring organophosphorus insecticide resistance have been selected at the *Rop-1* locus of the sheep blowfly, *Lucilia cuprina. Journal of Molecular Evolution* (in press).

60. Oakeshott, J.G., Devonshire, A.L., Claudianos, C., Sutherland, T.D., Horne, I., Campbell, P.M., Ollis D.L., and Russell, R.J. (2005). Comparing the organophosphorus and carbamate resistance mutations in cholin- and carboxyl-esterases. *Chemical and Biological Interactions* (in press).

61. Horne, I., Williams, M., Sutherland, T.D., Russell, R.J. and Oakeshott, J.G. A *Brevibacillus choshinensis* system that secretes cytoplasmic proteins. *Journal of Molecular Microbiology and Biotechnology* (in press).

Other Publications

62. Jackson, D.C., Brown, L.E., Russell, R.J., White, D.O., Dopheide, T.A. and Ward, C.W. (1980). Antigenic and immunogenic properties of influenza virus haemagglutinin fragments. In: Structure and Variation in Influenza Virus (ed. Laver/Air), pp. 309-320, Elsevier North Holland, Inc.

63. Russell, R.J., Dumancic, M.M., Foster, G.G., Weller G.L., Healy, M.J. and Oakeshott, J.G. (1990). Insecticide resistance as a model system for studying molecular evolution. In: Ecological and Evolutionary Genetics of Drosophila (Barker, J.S.F., Starmer, W.T. and MacIntyre, R.J. Eds.), pp. 293-314. Plenum, New York.

64. Oakeshott, J.G., Parker, A.G., Spackman, M.E., Cooke, P.H., Walker, V.K. and Russell, J.G. (1990). Molecular biological research in resistance management and alternative control strategies. In: Resistance Management in Parasites of Sheep (McKenzie, J.A., Martin, P.J. and Arundel, J.H. Eds.), pp. 53-69. Australian Wool Corporation, Melbourne.

65. Russell, R.J., Harcourt, R.L. and Oakeshott, J.G. (1998). Bioremediation of pesticides using enzymes. In ACIAR Proceedings No. 85: Seeking Agricultural Produce Free of Pesticide Residues. I.R. Kennedy, J.H. Skerritt, G.I. Johnson and E. Highley (eds), pp341-348.

66. Oakeshott, J.G., Claudianos, C., Campbell, P.M., Robin, G.C., Newcomb, R.D., Odgers, W.A., Harcourt, R.L. and Russell, R.J. (1998) Molecular basis of esterase-mediated OP resistance in two higher Diptera. In: Structure and Function of Cholinesterases and Related Proteins (Doctor BP, Taylor P, Quinn DM, Rotundo RL, Gentry MK., Eds.), pp. 497-502. Plenum Press, New York.

67. Oakeshott., J.G, Harcourt, R.L., Sutherland, T.D. and Russell R.J. (1999) Enzymatic

Bioremediation Technologies For Mycotoxins. In ACIAR Proceedings No 89: Elimination of aflatoxin contamination in peanut. R.G., Dietzgen, (ed.).

68. Russell, R.J., Harcourt, R., Sutherland, T., Nguyen, H. and Oakeshott, J.G. (1999) Can we clean water with microbes? Microbiology Australia. *20:* 14-17.

69. Claudianos, C., Crone, E. J., Coppin, C., Russell, R. J. and Oakeshott, J. G. (2001). A genomics perspective on mutant aliesterases and metabolic resistance to organophosphates. In: (Marshall Clark, J. and Yamaguchi, I. eds) *Agrochemical Resistance: Extent, Mechanism, and Dectection.* American Chemical Society, Washington. pp 90-101.

70. Sutherland, T.D., Russell, R.J., Nguyen, H. and Oakeshott, J.G. (2000) Bioremediation of endosulfan. Proceedings of the 10th Annual Cotton Conference, Brisbane, Australia.

71. EJ Crone, PM Campbell, TD Sutherland, RJ Russell and JG Oakeshott. (2000) Juvenile hormone esterase as transgenes for viral insecticides. Proceedings of the 10th Australian Cotton Conference. Brisbane, Australia.

72. Russell, R.J., Sutherland, T.D., Horne, I, Oakeshott, J.G., Zachariou, M., Nguyen, H.V., Selleck, M.L and Costello, M. (2001) Enzymatic bioremediation of chemical pesticides. Australasian Biotechnology *11:* 24-26.

73. Sutherland, TD, Horne, I, Weir, KM, Coppin, CW, Williams, MR, Briggs, LJ, Crone, EJ, Dorrian SJ, Russell, RJ, Selleck, ML, Costello, M, Nguyen, H, Dumsday, G, Zachariou, M, and Oakeshott, JG. (2002). Pesticide detoxification using enzymatic bioremediation. Proceedings of the 11th Australian Cotton Conference. Brisbane, Australia. pp 797-800

74. Crone, EJ, Sutherland, TD, Campbell, PM, Russell, RJ and Oakeshott, JG. (2002). Juvenile hormone esterase and three closely related esterases from *Drosophila melanogaster.* Proceedings of the 11th Australian Cotton Conference. Brisbane, Australia. pp 801-804

75. Oakeshott, JG, Claudianos, C, Campbell, PM, Newcomb, RD and Russell, RJ. (2005) Biochemical Genetics and Genomics of Insect Esterases. In: (Gilbert, L.I., Iatrou, K. and Gill, S. S. eds) *Comprehensive Molecular Insect Science* - vol. 5 - Pharmacology. Elsevier BV, Oxford UK. pp 309-381.

Patents:

1. Russell, R.J., Newcomb, R.D., Robin, C., Boyce, T.M., Campbell, P.C., Parker, A.G., Oakeshott, J.G. and Smyth, K-A. Enzyme Based Bioremediation. International Patent Application No PCT/AU95/00016 filed 13 January 1995. Priority claimed: 13 January 1994.

2.. Russell, R.J., Newcomb, R.D., Campbell, P.M., Robin, G.C., Claudianos, C., Smyth, K-A., Boyce, T.M., Oakeshott, J.G and Brownlie, J.C. Malathion Carboxylesterase. International Patent Application No PCT/AU96/00746 filed 23 November 1996. Priority claimed: 23 November 1995.

3. Sutherland, TD, Lacey, MJ, Russell, RJ and Oakeshott, JG. Improved endosulfan formulation and methods of use thereof. International Patent Application No PCT/AU01/01505 filed 20 November 2001. Priority date claimed: 20 November 2000.

4. Campbell, PM, Crone, EJ, Sutherland, TD, Russell, RJ and Oakeshott, JG. Juvenile hormone esterase. International Patent Application No. PCT/AU02/00101 filed 1 February 2002. Priority date claimed: 1 February 2001.

5. Russell, RJ, Heidari, R, Devonshire, A, Dorrian, SJ and Oakeshott, JG. Degradation of hydrophobic ester pesticides and toxins. International Patent Application No PCT/AU02/00114 filed 6 February 2002. Priority date claimed: 6 February 2001.

6. Oakeshott, JG, Devonshire, A, Coppin, CW, Heidari, R, Dorrian, SJ, and Russell, RJ Esterases with lipase activity. International Patent Application No PCT/AU02/00113 filed 6 February 2002. Priority date claimed: 6 February 2001.

7. Horne, I, Sutherland, T, Harcourt, R, Russell, R and Oakeshott, J. Phosphotriesterase from *Agrobacterium radiobacter* P230. International Patent Application No PCT/AU02/00594 filed 15 May 2002. Priority date claimed: 15 May 2001.

8. Sutherland, TD, Weir, K, Horne, I, Lacey, MJ, Russell, RJ and Oakeshott, JG. Enzymes which degrade thio compounds, and methods of use thereof. International Patent Application No PCT/AU03/00712 filed 6 June 2003. Priority date claimed: 6 June 2002.

9. Horne, I, Sutherland, T, Russell, R and Oakeshott, J. Expression system. International Patent Application No PCT/AU03/00951 filed 28 July 2003. Priority date claimed: 26 July 2002.

10. Horne, I., Sutherland, T.D., Russell, R.J., Oakeshott, J.G. and Khurana, J. Variants of phosphotriesterases with enhanced and/or altered substrate specificity. International Patent Application No PCT/AU2003/001675 filed 16 December 2003. Priority date claimed: 16 December 2003.

Genetica 90: 239–268, 1993.

Evolutionary genetics of *Drosophila* esterases

J.G. Oakeshott, E.A. van Papenrecht, T.M. Boyce, M.J. Healy & R.J. Russell
CSIRO Division of Entomology, GPO Box 1700, Canberra ACT 2601, Australia

Received and accepted 15 October 1993

Key words: Drosophila, esterases, multigene families, structure function models

Abstract

Over 30 carboxylester hydrolases have been identified in *D. melanogaster*. Most are classified as acetyl, carboxyl or cholinesterases. Sequence similarities among most of the carboxyl and all the cholinesterases so far characterised from *D. melanogaster* and other eukaryotes justify recognition of a carboxyl/cholinesterase multigene family. This family shows minimal sequence similarities with other esterases but crystallographic data for a few non-drosophilid enzymes show that the family shares a distinctive overall structure with some other carboxyl and aryl esterases, so they are all put in one superfamily of /β hydrolases. Fifteen esterase genes have been mapped in *D. melanogaster* and twelve are clustered at two chromosomal sites. The constitution of each cluster varies across *Drosophila* species but two carboxyl esterases in one cluster are sufficiently conserved that their homologues can be identified among enzymes conferring insecticide resistance in other Diptera. Sequence differences between two other esterases, the EST6 carboxyl esterase and acetylcholinesterase, have been interpreted against the consensus super-secondary structure for the carboxyl/cholinesterase multigene family; their sequence differences are widely dispersed across the structure and include substantial divergence in substrate binding sites and the active site gorge. This also applies when EST6 is compared across species where differences in its expression indicate a difference in function. However, comparisons within and among species where EST6 expression is conserved show that many aspects of the predicted super-secondary structure are tightly conserved. Two notable exceptions are a pair of polymorphisms in the substrate binding site of the enzyme in *D. melanogaster*. These polymorphisms are associated with differences in substrate interactions *in vitro* and demographic data indicate that the alternative forms are not selectively equivalent *in vivo*.

Introduction

It has been a cornerstone of evolutionary theory that the diversity of biochemical functions required in complex multicellular organisms has been achieved by the evolution of multigene families. Two types of component process are proposed. First, there must be an amplification of gene copy number through a set of duplication events. This, in turn, allows some of the duplicated genes to diverge from the ancestral gene and adopt new functions, without compromising the ability of other members of the duplicated family to carry out the ancestral function. The first process, the amplification of copy number, has been exemplified many times, through both classical genetic and molecular evidence for clusters of closely related genes. However, the subsequent process of functional divergence has proven less tractable because it requires detailed biochemical, genetic and molecular analyses.

One set of gene/enzyme systems that is now proving amenable to all three approaches is a subgroup of esterase enzymes that preferentially hydrolyse esters of carboxylic acids (E.C. 3.1.1). As is common usage we simply term this subgroup the esterases. It excludes a multitude of phospho- and thio-esterases, phosphatases and sulphatases but even so encompasses several multigene families. Much of the work to date has focussed on verte-

brate and to a lesser extent microbial esterases; sequence data are now available for many of these and crystal structures have also been solved for a few. However, a second strand of work has been exploiting the unique advantages of *D. melanogaster* to probe genetic and molecular aspects of the esterases that would be intractable with other higher organisms. The following review focusses on this second strand of work, but also seeks to reinterpret many of the findings in the light of the sequence-structure-function models developed from the crystal structures of the related enzymes in the first strand.

Functional classification of esterases

Mammalian and microbial precedents

While the substrate specificities of some esterases may be quite narrow *in vivo*, many hydrolyse a broad and overlapping range of substrates *in vitro*. One advantage of this has been that many can be assayed electrophoretically, by coupling their hydrolysis of various synthetic naphthyl esters to the colorimetric conversion of certain dyes. Over 20 esterase isozymes have been resolved by these means in each of several vertebrate species (e.g., Holmes & Masters, 1967; Coates, Mestriner & Hopkinson, 1975). However, the disadvantage of these broad substrate ranges is that other criteria are also required to achieve a functional classification of esterase activities. The classification most widely used relies primarily on sensitivities to diagnostic concentrations of three groups of inhibitors, namely sulfhydryl reagents (typically p-chloromercuribenzoic acid, or pCMB), organophosphates (OPs, such as paraoxon, fenitrooxon and diisopropyl fluorophosphate, or DFP) and eserine sulfate (Holmes & Masters, 1967). Four classes of enzymes are discriminated on these criteria:

Acetyl esterases, which are not affected by any of the inhibitors and generally prefer aliphatic substrates involving acetic acid.

Aryl esterases, which are only inhibited by the sulfhydryl reagents and generally prefer aromatic substrates.

Carboxyl esterases, which are only inhibited by the OPs and prefer aliphatic esters, generally of longer acids than acetic acid, and

Cholinesterases, which are inhibited by both OPs and eserine sulfate and prefer charged substrates like cholinesters over other aromatic or aliphatic esters.

At least half the esterases detectable by systematic electrophoretic surveys of mammalian species prove to be carboxyl esterases, most of the remainder comprising similar numbers of the other three classes (Ecobichon & Kalow, 1965; Holmes & Masters, 1967). However, a few mammalian esterases have now been described that do not fit readily into any of these four classes. One example from humans is butyrylesterase, which is inhibited by pCMB and OPs but not eserine sulfate (Hjoring & Svensmark, 1988).

Some but not all aspects of this mammalian-based classification system appear to hold for esterases in organisms as distantly related as prokaryotes. Enzymes classifiable as aryl and acetyl esterases against the inhibitor criteria have been described from bacteria (e.g., Choi *et al.*, 1990; Luthi *et al.*, 1990). However, we are unaware of any bacterial enzymes that would be classified as carboxyl or cholinesterases on these criteria. In fact, many bacterial esterases fall outside all four inhibitor-based classes (e.g., Raymer, Willard & Schottel, 1990, and references therein).

Drosophila *and other insects*

Twenty-two soluble esterase isozymes have been detected by native polyacrylamide gel electrophoresis (PAGE) of individual tissues and defined life stages of *D. melanogaster* (Healy, Dumancic & Oakeshott, 1991). At least seven more can be resolved if a second dimension of electrophoresis involving isoelectric focussing is applied (Healy, Dumancic & Oakeshott, 1991; Campbell, Healy & Oakeshott, 1992). Most of these esterases are probably encoded by different genes. The only exceptions found so far are four isozymes of acetylcholinesterase (AChE) that are all encoded by the one gene (*Ace*; Arpagaus, Fournier & Toutant, 1988, and see below). All five non-AChE isozymes which have been mapped genetically are inherited as single and distinct genes (see below) and this accords well with mammalian precedents. In mouse, for example, over 30 different non-AChE isozymes have been described (Ruddle & Harrington, 1967) and so far 23 have been shown to be products of different genes (O'Brien, 1990; Von

Deimling & Wassmer, 1991; Von Deimling & Gaa, 1992).

The electrophoretic survey of Healy, Dumancic and Oakeshott (1991) would not have detected some esterases in *D. melanogaster* that either cannot hydrolyse any of the naphthyl esters used (e.g., most lipases and malathion carboxyl esterase, or MCE; Smyth *et al.*, in preparation), or are tightly associated with membranes and require detergent treatment to solubilise (e.g., EST23; Spackman *et al.*, 1993). Add these to the 29 soluble enzymes detected by electrophoresis or isoelectric focussing and it seems likely that there will be well over 40 esterases in *D. melanogaster*.

Inhibitor analyses of the 22 soluble isozymes detected by native PAGE plus the membrane-bound EST23 reveals eleven carboxyl esterases, six cholinesterases and three acetyl esterases, with only three isozymes not clearly classifiable (Table 1). Moreover, the carboxyl and cholinesterases can each be divided further into subclasses on the basis of inhibition by OPs and pCMB. The five subclass I carboxyl esterases are qualitatively more sensitive to inhibition by OPs than the six subclass II isozymes, while the four subclass I cholinesterases (the four AChE isozymes) are much less sensitive to pCMB inhibition than are the two subclass II enzymes. The fact that no aryl esterases have yet been recovered from *D. melanogaster* does not necessarily indicate their absence in this species, since most of the mammalian aryl esterases characterised by Holmes and Masters (1967) have relatively high electrophoretic mobility and relatively low stability under heat or urea treatment. If the aryl esterases of *D. melanogaster* behave in a similar way they would probably not have been detected under the conditions used to survey this species (Healy, Dumancic & Oakeshott, 1991, and references therein). Two carboxylesterases (esterase 6, or EST6, and juvenile hormone esterase, or JHE; White, Mane & Richmond, 1988; Campbell, Healy & Oakeshott, 1992) and two cholinesterases (AChE and EST9; Fournier *et al.*, 1988; Morton & Singh, 1985) have been substantially purified from *D. melanogaster* and a few others have been similarly characterised from other drosophilid (see below) and non-drosophilid insects (e.g., Devonshire, 1977; Kao, Motoyama & Dauterman, 1985; Ziegler *et al.*, 1987; Mouchés *et al.*, 1987; Abdel-Aal *et al.*, 1988; Field *et al.*, 1993). Like their mammalian counterparts, almost all have a subunit size between 60 and 70 KDa, with a single active site per subunit. The only exception is a termite carboxyl esterase, which only has a subunit size of 40 KDa (Sreerama & Veerabhadrappa, 1991). The only other esterases as small as this are several aryl esterases and some others that cannot be classified against the inhibitor criteria (Gan *et al.*, 1991; Ollis *et al.*, 1992; Cygler, Schrag & Ergan, 1992; Cygler *et al.*, 1993). No insect aryl or acetyl esterase has yet been purified.

Table 1. Inhibitor classification of 23 esterase isozymes from *D. melanogaster* (from Healy, Dumancic & Oakeshott, 1991 and Spackman *et al.*, 1993).

Inhibitor-based class	Isozyme
Carboxyl esterase	
Subclass 1	EST12; EST15; EST16; EST18; EST23
Subclass 2	EST1; EST2; EST6; EST14; EST17; EST22
Cholinesterase	
Subclass 1	EST4; EST5; EST8; EST13*
Subclass 2	EST9; EST10
Acetyl esterases	EST19; EST20; EST21
Miscellaneous	EST3; EST7; EST11

* All four subclass I cholinesterases are isozymes of AChE.

Esterase gene families

The α/β hydrolase superfamily

Over 50 esterase genes have now been cloned, mainly from vertebrates and bacteria but also several from fungi and insects. Most of the bacterial enzymes and a few eukaryotic esterases that cannot be classified on the inhibitor criteria show no recognisable sequence similarity with the other systems (Ounissi & Courvalin, 1985; Markovic & Jornvall, 1986; Ray *et al.*, 1988; Choi *et al.*, 1990; Raymer, Willard & Schottel, 1990; Zschunke *et al.*, 1991), suggesting that enzymes with esterase activity may have multiple origins. However, most eukaryotic carboxyl and cholinesterases, including all the esterases so far cloned from insects and a few bacterial esterases that are either aryl esterases or unclassifiable on the inhibitor criteria, do show at

least some sequence similarity (Ollis *et al.*, 1992; Cygler, Schrag & Ergan, 1992; Cygler *et al.*, 1993; see also Okada & Wakabayashi, 1988; Farrell *et al.*, 1990; Van der Meer *et al.*, 1991). On this basis, a small proportion of the bacterial esterases but a substantial proportion of eukaryotic esterases would seem to have evolved from just one of the ancient lineages. For reasons that we now elaborate this lineage is called the α/β hydrolase superfamily.

Several enzymes in this lineage, including bacterial aryl esterases and eukaryotic carboxyl and cholinesterases, have been crystallised and their tertiary structures resolved (Cygler, Schrag & Ergan, 1992; Ollis *et al.*, 1992; Van Tilbeurgh *et al.*, 1993, for references), enabling sequence similarities to be related to structural similarities. Two diagnostic features of the primary sequence that contribute to the active site have been found to recur across these enzymes. Moreover, sequence comparisons and *in vitro* mutagenesis indicate that these features also occur in other members of the lineage that have not yet been crystallised (e.g., Di Persio, Fontaine & Hui, 1990, 1991; Di Persio & Hui, 1993, and references therein). No insect esterase has yet been crystallised but the sequences of all those whose genes have been cloned contain these features (Cygler *et al.*, 1993).

The first feature is a triad of non-contiguous residues, generally Ser-Asp-His, otherwise Ser-Glu-His or Cys-Asp-His (ordered according to their occurence in the primary sequence). This triad forms a charge relay that executes the hydrolytic reaction by donating a proton to the ester bond of the substrate. Significantly the same residues are involved in the catalytic triad found in many proteases, but the order of the residues in the primary sequence of all the esterases in this lineage differs from the proteases. The only esterases known to have a different order for their triad are the rodent 'granzymes' and their human 'serine esterase' homologues, which show a trypsin-like order for their catalytic triad and are generally placed in the serine protease family (Zschunke *et al.*, 1991, and references therein). Of the three alternatives found in the esterases in the α/β hydrolase superfamily, the Cys-Asp-His is found in the (bacterial) aryl esterases sequences, Ser-Glu-His in the eukaryotic cholinesterases and some of the eukaryotic carboxyl esterases, and Ser-Asp-His in a minority of the eukaryotic carboxyl esterases. This explains in part why OPs (which bind Ser) do not inhibit aryl esterases but do inhibit the others.

The second feature of primary sequence shared by esterases in the α/β hydrolase lineage is the

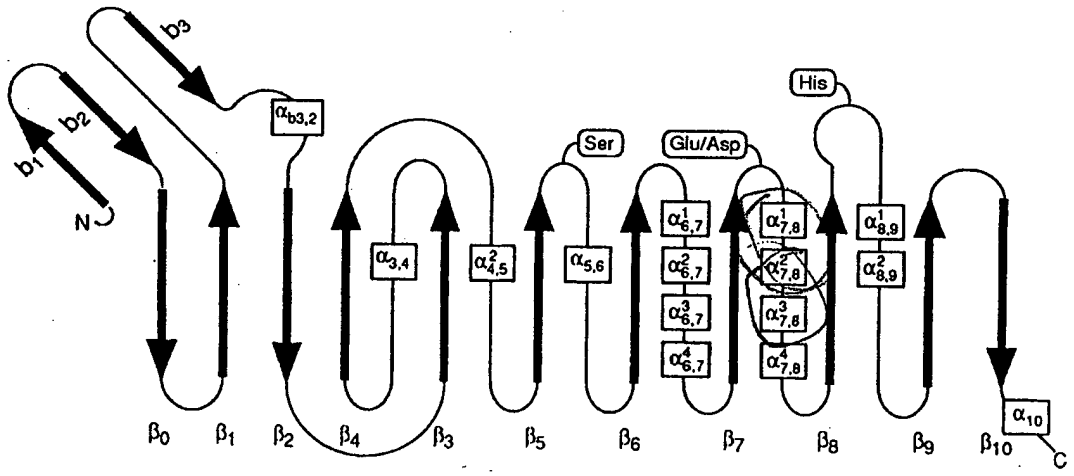


Fig. 1. Schematic diagram of the α/β hydrolase fold which Cygler *et al.* (1993) propose is shared by all members of the carboxyl/cholinesterase multigene family. β strands are shown as arrows and α helices are boxed. Following Cygler *et al.* (1993), β strands that contribute to the small and large b sheets are denoted β or β respectively, with their subscripted numbers indicating their order in the primary sequence. Subscripts for the α helices show the β strands they separate. The loops contributing to the catalytic triad are also shown. Coordinates for all structures in the primary sequences of AChE and EST6 are given in Table 1.

localisation of the Ser/Cys nucleophile of the catalytic triad in a highly conserved Gly-x-Ser/Cys-x-Gly pentapeptide. This motif is critical for the placement of the Ser/Cys at the apex of an extremely sharp turn and gives it ready access to the His of the triad on one side and the substrate on the other. This nucleophile pentapeptide, like the triad above, is also found in some protease families but, again excepting the granzyme/serine esterase group, its tertiary structure context is completely different, suggesting independent evolutionary origins.

Many of the esterases in the α/β hydrolase lineage show little sequence similarity with one another outside their catalytic triad and nucleophile pentapeptide. Yet the secondary and tertiary structures of five members of the group whose higher order structures have been resolved show substantial similarities with each other (Ollis *et al.*, 1992). Although their subunit sizes vary from about 25 to 60 KDa, they all show an internal backbone called the major β sheet, which comprises at least eight β strands. With a few exceptions, these strands show the same orientation with respect to each other and occur in the same order in the primary sequence. A minor β sheet of three strands and unknown function also occurs towards the amino terminus of the eukaryotic enzymes. Loops between the strands of the major β sheet include variable numbers of α helices and, amongst other functions, provide protection and stability to the major β sheet. The loops at the carboxy termini of (generally) strands β_5, β_7 and β_8 in the major β sheet juxtapose the triad residues, while excursions of variable size, mainly after strands β_6 and β_7, and before strand β_2, provide much of the substrate binding structure for the active site (see Fig. 1 and Table 2). Ollis *et al.* (1992) term this overall structure the α/β hydrolase fold and they argue that the extent of its conservation, even among enzymes with minimal primary sequence similarity, must reflect common, albeit ancient, ancestry.

The carboxyl/cholinesterase multigene family

Some members of the α/β hydrolase superfamily have non-esterase hydrolytic activities. However most of those characterised to date are esterases and these have been partitioned into six major families (Cygler, Schrag & Ergan, 1992). Between families

Table 2. The empirical relationship between primary and secondary structures for AChE from *T. california*, and the predicted relationship between these structures for EST6 from *D. melanogaster*. The AChE relationship and the nomenclature for secondary structures are taken from Sander and Schneider (1991) and Cygler *et al.* (1993). The EST6 relationship is based on the model of Cygler *et al.* (1993) and the alignment of EAvP. See also Figure 1 for a graphical representation of the secondary and supersecondary structures.

Primary sequence		Secondary structure	
AChE	EST6	Type	Name
7-10	8-11	β-Strand	b_1
13-16	14-17	β-Strand	b_2
18-21	19-22	β-Strand	β_0
26-34	23-31	β-Strand	β_1
57-59	55-57	β-Strand	b_3
–	–	α-Helix	$\alpha^1_{b3,2}$*
79-82	76-77	α-Helix	$\alpha_{b3,2}$
–	–	α-Helix	$\alpha^2_{b3,2}$*
96-101	86-91	β-Strand	β_2
109-115	100-106	β-Strand	β_3
133-139	122-128	α-Helix	$\alpha_{3,4}$
142-146	131-135	β-Strand	β_4
151-155	140-144	α-Helix	$\alpha^1_{4,5}$
168-183	156-171	α-Helix	$\alpha^2_{4,5}$
189-199	177-187	β-Strand	β_5
201-211	189-199	α-Helix	$\alpha_{5,6}$
221-225	209-213	β-Strand	β_6
238-251	224-238	α-Helix	$\alpha^1_{6,7}$
259-267	246-255	α-Helix	$\alpha^2_{6,7}$
271-277	258-265	α-Helix	$\alpha^3_{6,7}$
305-311	297-303	α-Helix	$\alpha^4_{6,7}$
319-324	311-316	β-Strand	β_7
329-335	321-327	α-Helix	$\alpha^1_{7,8}$
349-359	337-347	α-Helix	$\alpha^2_{7,8}$
365-376	353-363	α-Helix	$\alpha^3_{7,8}$
384-411	372,382-408†	α-Helix	$\alpha^4_{7,8}$
418-423	416-421	β-Strand	β_8
444-447	449-452	α-Helix	$\alpha^1_{8,9}$
460-479	464-483	α-Helix	$\alpha^2_{8,9}$
501-505	505-509	β-Strand	β_9
512-514	516-518	β-Strand	β_{10}
518-532	gap	α-Helix	α_{10}

* These helices are present in a closely related lipase from the fungus *Geotrichum candidum*, but absent in both AChE and EST6.

† $\alpha^4_{7,8}$ in EST6 is interrupted by a 10 residue insertion.

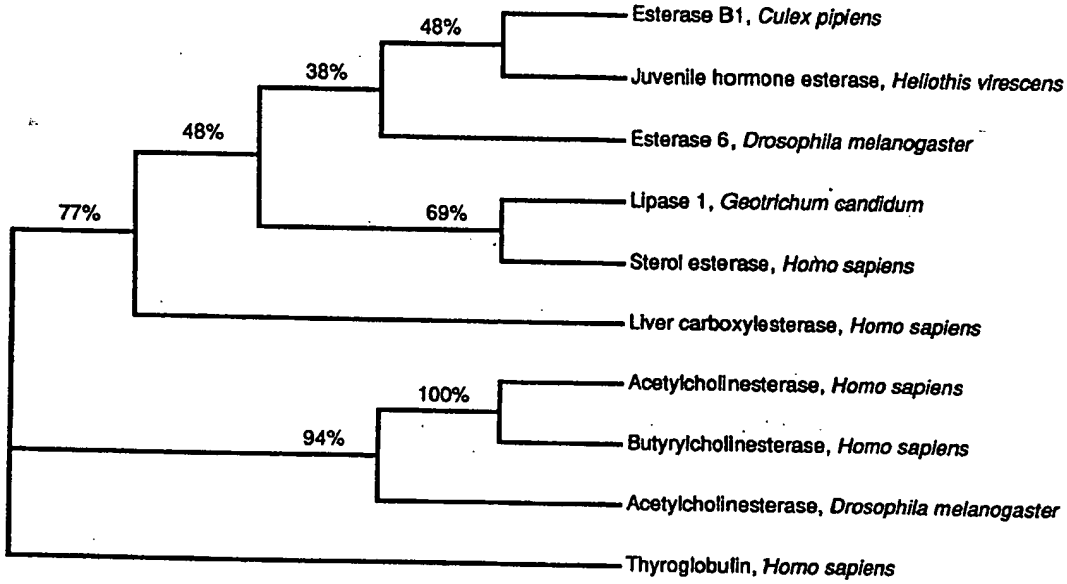


Fig. 2. Strict consensus network for the inferred amino acid sequences of ten members of the carboxyl/cholinesterase multigene family. Sequences were aligned as for Tables 1 and 2. Distances among sequences were calculated using the PAM distance matrix (Dayhoff, Schwartz & Orcutt, 1978) on one hundred randomly re-sampled (bootstrapped) replicates and networks were constructed using the neighbor joining method (Saitou & Nei, 1987) as implemented in PHYLIP version 3.4 (Felsenstein, 1991). The network is rooted with the human thyroglobulin sequence. Numbers to the left of each node indicate the percentage of replicates in which that node was formed. Line lengths are not proportional to distances.

primary sequence similarity is largely confined to the catalytic triad and nucleophile pentapeptide, but within families sequence similarities are much higher. One family contains several bacterial aryl esterases, the dienelactone hydrolases (Van der Meer *et al.*, 1991; Ollis *et al.*, 1992). Then there are four major families of lipases, two so far confined to bacteria and vertebrates, respectively, and two that include both fungal and vertebrate enzymes (plus at least three smaller families of lipases and cutinases). Finally, the largest and best characterised family includes some fungal lipases, some other eukaryotic carboxyl esterases and all the cholinesterases so far sequenced (Cygler, Schrag & Ergan, 1992, Cygler *et al.*, 1993). We term this latter family the carboxyl/cholinesterase multigene family. It covers all the insect esterases so far sequenced.

Sequences have been published for just over 30 carboxyl/cholinesterases. Seven are orthologous AChEs from different species, but most of the others are clearly paralogous. Across the whole family the average pairwise sequence identity is 29% and this figure rises to 36% for a stretch of just over 300 residues which lies in the amino terminal part of most of the proteins (see below for three exceptions). Higher order structures are known for two of the enzymes, a fungal lipase (Schrag & Cygler, 1993) and a vertebrate AChE (Sussman *et al.*, 1991; Rippol *et al.*, 1993). Although the functions and sequences of these two enzymes make them quite distant relatives within the family (see Fig. 2), they nevertheless show remarkable similarity in higher order structure. It has therefore been possible to predict with some confidence how the primary sequences of other family members contribute to secondary and super-secondary, if not tertiary, structure. For relatively closely related family members, like AChE and butyrylcholinesterase (BChE), where overall sequence identity exceeds 50%, the predictions can be extended to the point of identifying particular residues as candidates to explain very specific differences in substrate specificities. In some cases these predictions have been

validated by functional analyses of *in vitro* mutations of the candidate residues (e.g., Vellom *et al.*, 1993).

As with some of the other families of α/β hydrolases, the carboxyl/cholinesterase group includes a few proteins without esterase activity. In this case three family members have been identified that have no known enzymic activity and lack the catalytic triad and Gly-x-Ser/Cys-x-Gly pentapeptide. These exceptional proteins are the mammalian hormone precursor thyroglobulin and two other cell adhesion molecules, glutactin and neurotactin (Mercken *et al.*, 1985; Olson *et al.*, 1990; De la Escalera *et al.*, 1990). All three show high sequence similarity to other family members in the amino terminal 300 residue stretch that includes most of the essential scaffolding of β strands in the α/β hydrolase fold. While these β strands at least will be substantially confined to the protein interior, some loops between these strands must detour to the exterior to confer cell adhesion properties. This has been shown by functional analyses of various chimeric mutant proteins in which segments have been exchanged among glutactin, neurotactin and AChE (some forms of which are also associated with membranes) (Piovant *et al.*, 1993).

Just as there are some members of the carboxyl/cholinesterase family that have no known hydrolytic activity, there are also others which have two such activities, specifically amidase as well as esterase activities (Heymann, 1980; Richmond *et al.*, 1990, and references therein). Intriguingly, in the one case analysed mechanistically, a mammalian sterol esterase, the two activities appear to be at least partly independent (Hui, Hayakawa & Oizumi, 1993). Both activities are susceptible to OP inhibition, suggesting a key Ser residue in the active site, although it has not been formally established that it is the same Ser in both cases. However, the His in the catalytic triad required for esteratic activity can be mutated to Gln, abolishing esterase activity as expected but, surprisingly, leaving the amidase activity largely unaltered.

The carboxyl/cholinesterase split

Figure 2 shows a neighbor joining network (Saitou & Nei, 1987; Felsenstein, 1985) relating the inferred amino acid sequences of nine members of the carboxyl/cholinesterase multigene family. Four of them are from insects, four from humans and one from a fungus. Apart from the *Drosophila* and human AChEs, the network only includes genes that are clearly paralogous and the functions of which are known, at least in part. Human thyroglobulin is also included as a non-enzymic member of the family to provide an outgroup for the network.

Three major splits are identified in the network. There is a primary separation of carboxyl and cholinesterases and then a further bifurcation within each of these two lineages. Within the former the lipases are separated from other carboxylesterases and within the latter the human representatives are separated from the insect enzyme. Thus the functional difference between carboxyl and cholinesterases is reflected in their sequence divergence, regardless of the taxa from which they originate. Furthermore, since both the carboxyl and cholinesterase lineages include insect and human genes, the carboxyl/cholinesterase split is likely to have preceded the divergence of vertebrates and invertebrates. Indeed, it could be even older, since the lipase lineage within the carboxyl esterases includes both a fungal and a vertebrate enzyme. On the other hand, within the cholinesterases human AChE and human BChE share a more recent common ancestry than either do with the insect AChE, which suggests a relatively recent acquisition of BChE function within the vertebrate line. Consistent with this relationship, no enzyme with obvious biochemical homology to BChE has been described from *Drosophila*, or any other insect.

Esterase genetics in *D. melanogaster*

The previous sections have established a hierarchy of evolutionary relationships among various functional classes of esterases. In order to investigate the processes that have given rise to this pattern we must first examine the genomic organisation of the cognate genes. The power of *D. melanogaster* genetics suits it well for this purpose. Genes for a total of eight esterase activities have now been mapped in this species by classical genetics. These include three carboxyl and two cholinesterases and three whose status against the inhibitor-based criteria has not been determined. Several other putative esterase genes have also been discovered during the course of sequence analyses of cloned DNA. Col-

lating the classical and molecular genetics indicates a minimum of 15 genes at five separate chromosomal sites, one of which contains two genes and the other at least ten (Fig. 3).

The three genes so far attributed to unique sites produce enzymes with very diverse functions. These are:

Ace, which encodes the AChE isozymes, or subclass I cholinesterases. Their preferred naphthyl ester substrate *in vitro* is a naphthyl acetate and their *in vivo* substrate is the neurotransmitter acetylcholine (Fournier *et al.*, 1988, and references therein). Hybridisation of *Ace* clones to various chromosomal deficiencies localises it to 87E3 on the polytene map (3-52 genetically) of chromosome IIIR (Spierer *et al.*, 1983).

Est17, which encodes the EST17 isozyme, a subclass II carboxyl esterase whose preferred naphthyl ester substrate *in vitro* is α naphthyl proprionate. The *in vivo* function of EST17 is unknown but it is largely confined to late larvae (Healy, Dumancic & Oakeshott, 1991). Mapping of allozymic variants against the *rusteca* chromosome in 107 test cross progeny places *Est17* roughly equidistant from *e* (3-70.7 genetically) and *ca* (3-100.7), which puts it distal to *Ace* on chromosome IIIR (P. Kostakos, R.J.R. & J.G.O., unpublished data).

Est9 of Loukas (1981). The EST9 isozyme encoded by this gene hydrolyses α naphthyl acetate *in vitro*, but only in the presence of the peptidase substrate l-leucyl β-naphthyl amide. The latter may have some mechanistic analogy to the joint esterase-amidase activities of the mammalian sterol esterase (see above). However, in the absence of other inhibitor analyses, the status of this EST9 is uncertain against the inhibitor criteria. Mapping of allozymic variants puts the *Est9* of Loukas (1981) on chromosome II and, assuming homology with a similar enzyme in *D. pseudoobscura*, probably on IIR (Loukas, 1981).

An enzyme product has so far only been identified for one of the two genes in the smaller of the

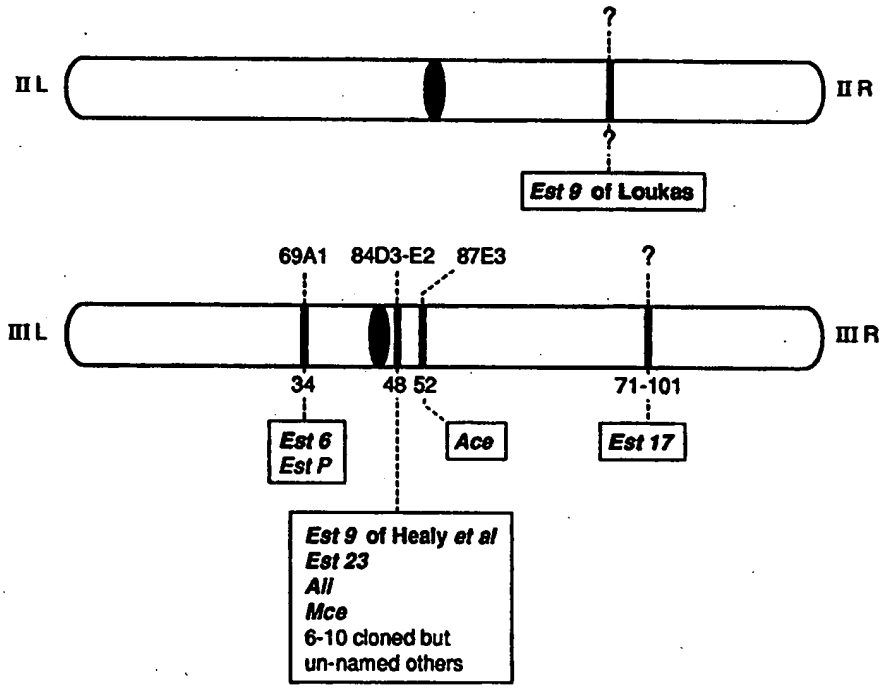


Fig. 3. Chromosomal locations of *D. melanogaster* esterase genes. Bars illustrate chromosomes II and III, with the left (L) and right (R) arms indicated. Cytological locations are indicated above the chromosomes and genetic map positions below. *Est9* of Loukas (1981) has not been localised within chromosome IIR. Genes with known biochemical phenotypes are in bold; cloned genes for which no biochemical phenotypes have yet been identified are in normal typeface.

two esterase gene clusters. This pair includes:
Est6, which encodes the EST6 isozyme, an abundant and widely distributed subclass II carboxyl esterase whose preferred naphthyl ester substrate *in vitro* is β naphthyl acetate. Its *in vivo* substrate is unknown but physiological and behavioural data indicate a role in reproductive fitness (see below). *In situ* hybridisation of cloned *Est6* to various chromosomal deficiencies localises *Est6* to 69A1 on the polytene map (about 3-34 on the genetic map) of chromosome IIIL (Oakeshott *et al.*, 1987; Procunier, Smith & Richmond, 1991).

EstP, which is downstream from *Est6* at 69A1. It lies in the same orientation as *Est6* and its initiation codon is just under 200 bp 3′ of the *Est6* termination codon (Collet *et al.*, 1990). It is mainly transcribed in late larvae but bears no obvious correspondence with the developmental profile of any isozyme detected by Healy, Dumancic and Oakeshott (1991).

Four biochemically defined esterase activities have so far been mapped to the larger cluster of ten esterase genes. These four are functionally diverse, in that they include both carboxyl and cholinesterases on the inhibitor criteria. However, a common feature of three of the activities at least is their homologies to various esterases implicated in OP insecticide resistance in other insects. The genes for these four activities are so tightly clustered that they have not been separable by deficiency mapping. They are:
Est9 of Healy, Dumancic and Oakeshott (1991) (= *EstC* of Beckman and Johnson, 1964), which codes for the abundant subclass II cholinesterase isozyme EST9 (ESTC). The preferred naphthyl substrates for this enzyme *in vitro* are α naphthyl proprionate and butyrate. Its *in vivo* substrate is unknown but the high concentration of the enzyme in the gut suggests a role in the digestion of dietary esters (see below). Localisation of the gene to 84D3-5 (about 3-48 genetically) on chromosome IIIR was achieved by deficiency mapping of EST9 allozyme variants (Cavener, Otteson & Kaufman, 1986).

Ali, which has not been related to a specific esterase isozyme but which accounts for the majority of *in vitro* methyl butyrate hydrolysis in a manometric assay. The *in vivo* function of ALI is unknown but it shows strong biochemical similarity to the ALI activity which is greatly reduced in OP resistant mutants of the house fly *Musca domestica* (Oppenoorth & Van Asperen, 1960). The co-localisation of *Ali* with *Est9* above involved assays of ALI activities in deficiency heterozygotes (Spackman *et al.*, 1993).

Est23, which encodes the microsomal subclass I carboxyl esterase isozyme EST23. The preferred naphthyl ester substrate of this isozyme *in vitro* is α naphthyl acetate. The *in vivo* function of EST23 is unknown but the enzyme shows very close biochemical similarity with the E3 enzyme which mutates to an apparently null phenotype in OP resistant mutants of the Australian sheep blowfly *Lucilia cuprina* (Spackman *et al.*, 1993). Co-localisation of *Est23* with *Ali* and *Est9* was achieved by deficiency mapping of EST23 allozymes (Spackman *et al.*, 1993).

Mce, which has not been related to any specific esterase isozyme but which encodes the ability to hydrolyse a small subset of OPs like malathion which carry carboxyl ester linkages in their alcohol group (in addition to the phosphoester linkage common to all OPs). Its *in vivo* function is unknown but elevated MCE levels have been implicated in malathion resistance in several pest species (Ziegler *et al.*, 1987; Russell *et al.*, 1990). Co-localisation with *Est9*, *Ali* and *Est23* involved assays of MCE activities in deficiency heterozygotes (Spackman *et al.*, 1993).

Despite the differences in their biochemical phenotypes, the precise genetical relationships among *Est9*, *Ali*, *Est23* and *Mce* are as yet uncertain. It is clear that *Est9* and *Est23* are each single and distinct genes, because the allozymic variants they encode vary independently across strains (Spackman *et al.*, 1993). There is also indirect evidence that *Est23* is distinct from *Mce*, in that their homologues in *L. cuprina* are separable by recombination (albeit at very low frequencies; Raftos & Hughes, 1986; Smyth *et al.*, in preparation). However, in the absence of electrophoretic phenotypes, we cannot discount the possibility that the spectrophotometrically or radiometrically determined ALI and MCE phenotypes are each due to the products of more than one closely linked genes. For ALI in particular, it is possible that such genes include either or both of *Est9* and *Est23*.

On the other hand, molecular data show that it is at least possible for all four of the biochemical phenotypes to be encoded by distinct genes. A total of ten putative esterase genes have been identified

from molecular studies of a 90 kb stretch of contiguous DNA at 84D3-E2 (R.J.R., C. Robin, P. Kostakos, R. Newcomb, L. Court, K. Medveczky, D. Hartl, T.B. & J.G.O. unpublished data). Preliminary analyses indicate that at least eight of the ten genes may be active and that they encompass part, if not all of the *Est9-Ali-Est23-Mce* cluster identified by the classical genetics. However, the precise correspondence of individual genes detected by the two approaches has not yet been elucidated.

Whatever the details of its molecular basis, the clustering of the *Est9-Ali-Est23-Mce* genes is intriguing in the context of the evolution of the carboxyl/cholinesterase multigene family, because EST9 is a subclass II cholinesterase and EST23 a subclass I carboxyl esterase. Two interpretations seem possible, given the phylogenetic distinction between at least some carboxyl and cholinesterases suggested by the neighbor joining network in Figure 2. One interpretation is that the cluster is ancient and pre-dates the divergence of carboxyl and cholinesterases. Alternatively, there may have been more than one carboxyl/cholinesterase split within the family, perhaps one for the subclass I cholinesterases revealed by the network in Figure 2 and one for subclass II cholinesterases represented by the *Est9-Ali-Est23-Mce* cluster.

Evolution of esterase gene families across *Drosophila* species

Having found evidence for two clusters of esterase genes in *D. melanogaster*, we now take a comparative approach to examine the evolution of the two clusters across *Drosophila* and other insect species. While not studied in the same detail as *D. melanogaster*, several other *Drosophila* species have also been surveyed for soluble esterases by electrophoretic analysis, generally of whole fly homogenates (e.g., Johnson *et al.*, 1966; Johnson, Richardson & Kambysellis, 1968; Berger & Canter, 1973; Mulley, James & Barker, 1979; Baker, 1980; Morton & Singh, 1985; Korochkin *et al.*, 1987). As many as ten isozymes have been detected in some species, but a recurrent finding is the presence of two intensely staining non-AChE isozymes in extracts from a variety of life stages.

In *D. melanogaster* these isozymes are EST6 from the small gene cluster above and the EST9 of Healy, Dumancic and Oakeshott (1991) from the larger cluster. However, the corresponding isozymes take other names in many of the other species. Generally, the two isozymes have been simply distinguished by their *in vitro* preference for α versus β naphthyl ester substrates (which gives rise to differently coloured bands when coupled with several histochemical dyes). This α/β preference itself does not necessarily indicate a large biochemical difference; another minor isozyme of *D. mojavensis* is actually polymorphic for α- and β-preferring forms (Zouros & Van Delden, 1982). In the case of the two major isozymes however, it is consistently associated throughout the genus with a number of other distinct biochemical and physiological properties.

The α-esterase cluster

The major α-preferring isozyme has been characterised biochemically from members of the melanogaster and obscura groups (mainly *D. melanogaster* and *D. pseudoobscura*, respectively) in the subgenus Sophophora and the virilis (mainly *D. virilis*) and repleta groups (*D. mojavensis*, *D. buzzatii*) in the subgenus Drosophila (Narise, 1973; Sasaki & Narise, 1978; East, 1982; Morton & Singh, 1985).

The enzyme has been at least partially purified (over 150-fold in some cases) by various size fractionating chromatographic procedures in all four lineages and proves to be a monomer of 55-70 KDa in all cases. In crude homogenates the enzyme from representatives of each lineage is highly sensitive to OPs and somewhat sensitive to sulfhydryl reagents and eserine sulfate, classifying it as a subclass II cholinesterase on the inhibitor criteria (Morton & Singh, 1985; Healy, Dumancic & Oakeshott, 1991; M.M. Dumancic, M.J.H. & J.G.O., unpublished data). One distinctive property of the enzyme in all four lineages is that it is strongly stabilised by β-mercaptoethanol, suggesting that its sensitivity to sulfhydryl reagents represents an unstable disulfide bond (Morton & Singh, 1985; M.M. Dumancic, M.J.H. & J.G.O., unpublished data). Another shared property is a widespread distribution across tissues but major concentration in gut tissue throughout the life cycle (Kambysellis, Johnson & Richardson, 1968; Sasaki, 1974; East, 1982; Morton & Singh, 1985). The enzyme is also weakly

associated with membranes in several of the species (Korochkin, Matveeva & Kerkis, 1973; M.M. Dumancic, M.J.H. & J.G.O., unpublished data). Thus, in terms of its biochemistry and physiology at least, there is a good case for the orthology of the major α-preferring isozymes across the four lineages.

Although not inconsistent with this conclusion, the one immunological study of the enzyme to date does show substantial epitope divergence among the species groups. Thus, Saski (1975) found that polyclonal antibody to the major α-preferring enzyme of *D. virilis* cross-reacted with other species within the virilis group, but not outside it.

Consistent with the biochemistry and physiology, classical genetic analyses of allozyme variants also implies orthology of the major α-preferring isozyme across the four species groups. The structural gene *Est9* maps to chromosome IIIR in *D. melanogaster* (see above) and the structural genes for the major α-esterases in the other species analysed map to the homologous chromosomes/arms (Triantaphyllidis & Christodoulou, 1973; Tsuno, Aotsuka & Ohba, 1984; Morton & Singh, 1985; Schafer *et al.* 1993).

We have shown in the previous section that *Est9* of *D. melanogaster* lies in a cluster of about ten esterase genes. The three other enzyme activities so far traced to this cluster, ALI, EST23 and MCE, are significantly different from EST9 and its putative orthologues in the other species in terms of their electrophoretic mobilities and substrate and inhibitor specificities. However, where information is available, they do resemble EST9 in being α-esterases associated with gut tissues and membranes. Likewise, there is evidence for additional α-esterase genes clustered around the putative *Est9* orthologue in the virilis group of species (Baker, 1980; Korochkin *et al.*, 1987). None of these enzymes were tested for ALI or MCE activities but one appears similar to EST23. On the basis of the virilis data, Korochkin *et al.* (1987) proposed the existence of an α-esterase gene cluster. We suggest that it is orthologous to the *Est9-Ali-Est23-Mce* cluster in *D. melanogaster*. The constitution of the two clusters may not be identical but multiple members of both clusters are membrane-associated gut α-esterases.

There is also evidence for such a cluster in another dipteran, *L. cuprina*. The E8 isozyme of *L. cuprina* shows strong similarities to EST9 of *D. melanogaster* both *in vitro*, in terms of its electrophoretic phenotype, and *in vivo*, in its developmental profile and tissue localisation (Parker *et al.*, 1991). E8 is monomorphic electrophoretically, so has not been mapped genetically. However, three other esterases have been mapped to a 1cM region on the *L. cuprina* homologue of chromosome IIIR that contains the *Drosophila* α-esterase cluster (Raftos & Hughes, 1986; Smyth *et al.*, in preparation; see also Lai-Fook & Smith, 1991). Two of these enzymes, E3 and MCE, are clearly homologous on biochemical and physiological criteria to EST23 and MCE within the *D. melanogaster* cluster (Russell *et al.*, 1990; Parker *et al.*, 1991; Spackman *et al.*, 1993; Smyth *et al.*, in preparation). The other enzyme, E9, has no obvious homologue in *D. melanogaster* but it is interesting that it is another microsomal α-esterase (Parker *et al.*, 1991; Smyth *et al.*, in preparation).

We predict that the *E3, E8, E9* and *Mce* genes will be in the *L. cuprina* homologue of the array of ten cloned genes to which we ascribe the α-esterase cluster of *D. melanogaster*. The high level of conservation of some α cluster genes implied by the *L. cuprina* comparisons is also evident in a neighbor joining network of all insect carboxylesterases for which full sequence data are available. This network (not shown) indicates a close relationship between enzymes conferring OP insecticide resistance on mosquitoes and the EST9-EST23-ALI-MCE cluster in *D. melanogaster*, suggesting that esterases involved in OP resistance may be homologous across several Diptera. This network also shows a clear separation between the α and the β cluster of *D. melanogaster* described below.

The β-esterase cluster

As with the α-preferring isozyme, the major β-preferring esterase has been characterised biochemically from representatives of the virilis, repleta, melanogaster and obscura species groups (Narise & Hubby, 1966; Narise, 1973; Sasaki & Narise, 1978; Mane, Tepper & Richmond, 1983; East, 1984; Pen, Rongen & Beintema, 1984; Pen, Van Beeumen & Beintema, 1986; Morton & Singh, 1985; White, Mane & Richmond, 1988; Farmer & Carter, 1989). For each group this includes purification, at least

partially, and in some cases to homogeneity (entailing about 200-300 fold purification, depending on the species).

The major β-esterase is a homodimer variously estimated at 100-140 KDa in the great majority of species investigated. However, it is a monomer of half the molecular weight in three species, including *D. melanogaster*, within the melanogaster complex of the melanogaster group (Morton & Singh, 1985). That the two forms of the enzyme are nevertheless orthologous is supported by the cross-reactivity of antibodies raised against the enzyme from *D. melanogaster* with the dimeric enzyme found in some of the other species (D. Morris & R.C. Richmond, cited in Oakeshott, Healy & Game, 1990). Consistent with this, *D. pseudoobscura*, in the obscura group, usually has a dimeric form of the enzyme but monomeric variants also occur (Arnason & Chambers, 1987). Across all species studied the two forms of the enzyme also have qualitatively similar inhibitor sensitivities, being sensitive to eserine sulfate and OPs, but not sulfhydryl reagents. Although polyclonal antibodies against the enzyme from virilis, repleta and melanogaster species all show wide cross-reactivity with the other species groups they do not cross-react with their major α-preferring enzymes, and in the case of virilis the converse also holds (Sasaki, 1975; Pen, Van Beeumen & Beintema, 1986; D. Morris & R.C. Richmond, cited in Oakeshott, Healy & Game, 1990). Thus, all available biochemical and immunological evidence suggests that the major β-preferring isozymes in the different lineages are orthologous to each other. These data also support the neighbor joining network above in indicating a clear distinction from the major α-preferring enzymes.

Unlike the α-enzymes, some clear evolutionary differences emerge in a comparison of the stage and tissue distributions of the major β-preferring enzymes across species. The enzyme is found in the hemolymph of all the species so far investigated but some additional sites of expression prove to be phylogenetically restricted. Two examples are male reproductive tract expression in some of the melanogaster species complex and eye expression in *D. pseudoobscura* (Morton & Singh, 1985; Brady & Richmond, 1990; Oakeshott, Healy & Game, 1990).

Two further complications for the orthology argument arise from a comparison of the genetics of the major β-preferring enzyme across species. First, the structural gene is loosely linked to the α-esterase cluster and on the same chromosome arm in all the virilis and repleta species analysed, but lies on a different arm in the melanogaster and obscura species studied (IIIL in *D. melanogaster*; Morton & Singh, 1985; Korochkin *et al.*, 1987; Oakeshott, Healy & Game, 1990; Schafer *et al.*, 1993). Second, while the gene lies in another cluster of esterase sequences in all four lineages, the organisation of that cluster appears to vary substantially among the lineages. Current knowledge of the cluster, termed the β-esterase cluster by Korochkin *et al.* (1987), can be summarized as follows.

For the melanogaster group: As outlined in the previous section, EST6 in *D. melanogaster* is expressed in both hemolymph and male reproductive tract. Molecular analysis shows the *Est6* gene to be the 5′ member of a tandem pair of esterase sequences separated by about 200 bp (Collet *et al.*, 1990). Amino acid sequences inferred for the products of the two genes are 59% identical. No electrophoretic phenotype has been identified for the 3′ member of the pair (*EstP*) but transcript analysis shows its expression to be largely confined to late larvae.

For the obscura group: A 12 kb genomic fragment containing three esterase sequences has been isolated from *D. pseudoobscura* using *D. melanogaster Est6* as a probe (Brady, Richmond & Oakeshott, 1990; Brady & Richmond, 1990, 1992). Amino acid identities inferred for the three presumptive gene products lie between 65% and 81% and their identities with the melanogaster pair lie between 61% and 70%. The central gene is orthologous to *Est6* and encodes the major hemolymph β-esterase. Products and expression patterns have not been determined for the two flanking genes.

For the virilis group: Classical genetic analysis of *D. virilis* shows a cluster of three β-esterase genes, encoding the major hemolymph enzyme, a slower migrating form mainly expressed in male reproductive tract and a faster form mainly expressed in pupae (Korochkin *et al.*, 1987). A 15 kb clone proposed to contain two of these genes has also been isolated (Enikolopov *et al.*, 1989). Sequence analysis puts the two presumptive esterase genes in this clone about 500 bp apart, while genomic Southern analysis suggests the existence

of a third, closely related gene somewhere outside this 15 kb. Inferred gene products for the two cloned genes are about 50% identical to the melanogaster and obscura genes above. Curiously, both inferred gene products deviate significantly from other carboxyl/cholinesterases at several sites that are otherwise highly conserved in this family (Sergeev *et al.*, 1993; M.Z. Ludwig, N.A. Tamarina & R.C. Richmond, pers. comm.). Also, their expression patterns are very different from those of *Est6*/*EstP* in *D. melanogaster*. The 5′ member of the cloned *D. virilis* pair is known to encode the male reproductive tract enzyme and the 3′ member is proposed to encode the hemolymph enzyme (Korochkin *et al.*, 1990; Sergeev *et al.*, 1993).

For the repleta group: Classical genetic analysis of *D. mojavensis* shows a cluster of two genes, one encoding the hemolymph β-esterase and the other producing a slightly slower migrating isozyme in late larvae which is polymorphic for α- and β-preferring forms (Zouros *et al.*, 1982; Zouros & Van Delden, 1982). The N-terminal 34 residues of the two purified proteins are 82% identical and there is also strong similarity in their overall amino acid compositions (Pen, Van Beeumen & Beintema, 1986). Several repleta group species also show a separate, slower migrating β-esterase isozyme in male reproductive tract (Kambysellis, Johnson & Richardson, 1968) but the cognate gene has not yet been mapped. Using *D. melanogaster Est6* as a probe, East, Graham and Whitington (1990) cloned a 16 kb genomic fragment from *D. buzzatii* that may contain at least two and probably three esterase or esterase-like genes; sequence analysis of two of these revealed high similarity to the other esterase genes above (66% amino acid identity with each other, 47-50% with the melanogaster genes). Intriguingly, however, both the *D. buzzatii* sequences encode a Gly in place of the active site Ser, a change which would almost certainly render the proteins catalytically inactive. Since *D. buzzatii* does, in fact, produce active hemolymph and late larval β-esterase isozymes, there may be two clusters of β-esterase or β-esterase-like sequences in this species.

It is clear then that many aspects of the β-esterase cluster are highly variable and rapidly evolving, even within the genus *Drosophila*. Three properties which do recur across the lineages examined are expression in hemolymph, male reproductive tract and late larvae. However, in the subgenus Sophophora the first two of these properties are apparently expressed by the one enzyme whereas in the subgenus Drosophila they are expressed by different enzymes. Additionally, there are differences in the number of genes in the cluster, differences in the substrate specificities of the late larval enzyme (at least *in vitro*), the accumulation of sequence motifs in the *D. virilis* enzymes that are otherwise unusual in the carboxyl/cholinesterase multigene family, and the intriguing possibility of non-catalytic functions for the sequenced *D. buzzatii* genes.

Molecular bases for esterase evolution: cross-functional comparisons

Having identified patterns of conservation and change in the evolution of esterase functions across genes and species in the last two sections, we now examine their molecular bases. To assess structural changes we use the generic structure-function model for the carboxyl/cholinesterase multigene family outlined earlier. Where appropriate we also call on our knowledge of promoter functions to explain regulatory changes. We begin with two comparisons involving qualitative differences in esterase function, to be followed by two involving qualitatively conserved functions.

EST6 from D. melanogaster *versus AChE*

Figure 4 and Table 3 (top row) show how the generic model of Cygler *et al.* (1993) for the supersecondary structures of carboxyl/cholinesterases applies in the specific case of EST6 from *D. melanogaster*. The model allows specific supersecondary structures to be ascribed to about half of the EST6 primary sequence. The particular structural features that we are able to identify are the active site gorge, substrate binding sites, salt and cysteine bridges, the two β sheets and the various α helices (see Table 4 for precise coordinates). The model also suggests that many of the regions to which no specific structural feature could be ascribed lie in hydrophilic regions on the protein surface; such hydrophilic/surface regions have proved to be amongst the most variable parts of many other proteins (Chothia & Lesk, 1986).

Figure 4 and Table 3 also relate the predicted

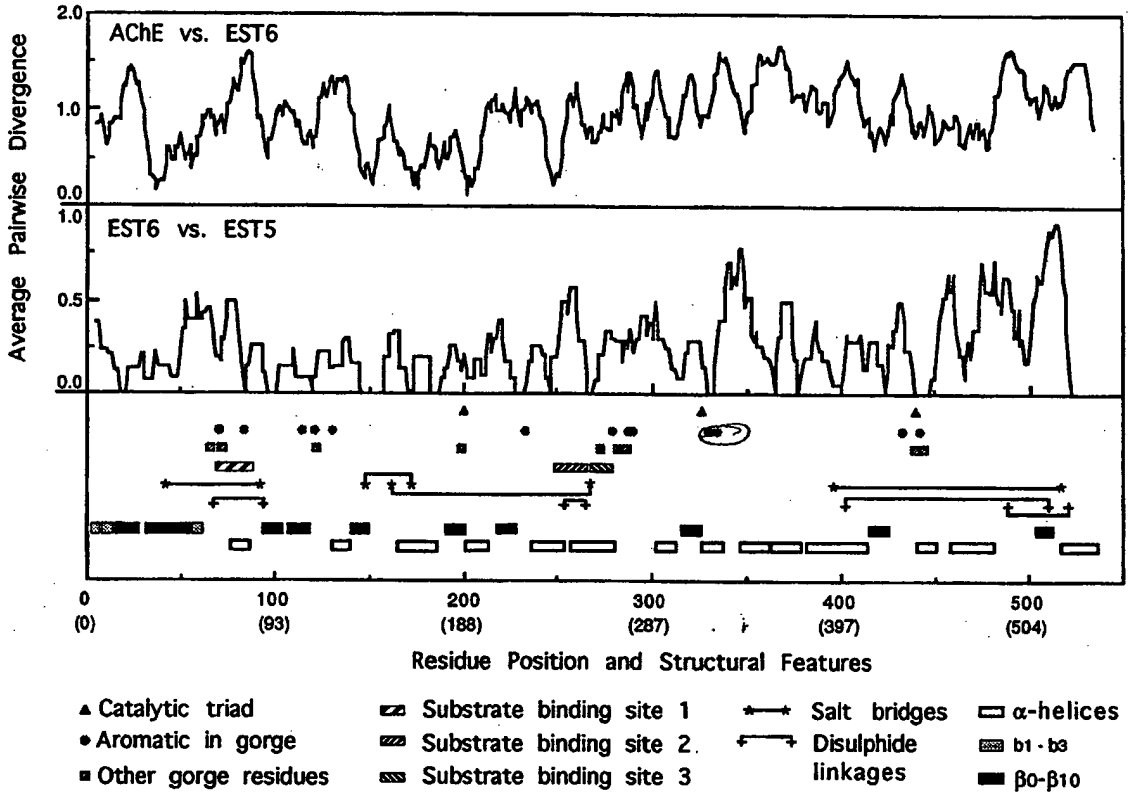


Fig. 4. Sliding window plots of the average pairwise distances between the mature protein sequences for EST6 from *D. melanogaster* and AChE from the electric eel, *Torpedo californica* (above) and EST6 and EST5 from *D. pseudoobscura* (below). Distances are calculated using the Dayhoff probability matrix (Dayhoff, Schwartz & Orcutt, 1978) normalised to give a mean of 1.5 and variance of 1.0. The plot was generated by calculating average distances for windows of nine residues moved along the sequence one residue at a time. Values are plotted above the central residue in each window and the AChE and EST6 (in parentheses) coordinates for the central residues are given on the abscissa. Below this axis are shown the various consensus structural elements which Cygler *et al.* (1993) identified for carboxyl/cholinesterases (see also Fig. 1 and Tables 2 and 4).

structure for EST6 to the sequence differences between EST6 and *T. californica* AChE, which is one of the crystallised enzymes on which the structural model is based. Sequence differences between EST6 and AChE are widespread across all the regions distinguished. The features showing least variation in Table 3 are the minor β sheet, the various salt andcysteine bridges, and the catalytic triad at the base of the active site gorge. The function of the minor β sheet is unclear but relatively strong conservation might be expected for the other two elements, because they will be critical for the maintenance of the α/β hydrolase fold structure (notwithstanding the relatively high variation in the major β sheet itself) and for esteratic activity. One notable difference involves the acid residue in the catalytic triad which is a Glu in AChE, but Asp in EST6. Some of the conserved elements identified in Table 3 correspond to troughs in the graphical plot of divergence against primary sequence in Figure 4, although the averaging approach in the latter tends to obscure conservation at the level of individual residues.

Two of the most divergent features identified in Table 3 are within the active site gorge and in substrate binding regions. Variation in the gorge is particularly concentrated in the so-called aromatic guidance residues which are involved in the movement of substrate down the AChE gorge (Rippol *et al.*, 1993). AChE contains 14 of these residues, but EST6 con-

Table 3. Amino acid sequence divergence among structural regions of AChE from *Torpedo californica* and EST6 from various *Drosophila* species and strains. Divergence is estimated as $\hat{d}$ (equation 4.5 of Nei, 1987) and its standard error from the variance of $\hat{d}$ (equation 4.6 of Nei, 1987) is in parentheses. Structural regions are defined in Table 4. Categories are exclusive so no amino acid position is used more than once.

	Gorge residues				Substrate binding regions				Salt, cysteine bridges	Sheets			
	catalytic	aromatic	other	Total	Site 1	Site 2	Site 3	Total		b_{1-3}	β_{0-10}	α-helices	Remaining residues
Number of residues	3	14	11	28	11	15	8	34	16	13	68	134	226
D. melanogaster EST6 vs. *T. californica* AChE	0.41 (0.41)	2.64 (0.96)	1.70 (0.64)	1.72 (0.41)	1.30 (0.49)	1.32 (0.43)	2.08 (0.94)	1.45 (0.31)	0.58 (0.22)	0.77 (0.30)	1.45 (0.22)	1.40 (0.15)	1.36 (0.11)
D. melanogaster vs. *D. pseudoobscura* EST6	0	0.24 (0.14)	0.20 (0.14)	0.20 (0.09)	0.45 (0.23)	0.41 (0.18)	0.29 (0.20)	0.39 (0.12)	0.06 (0.06)	0.26 (0.15)	0.25 (0.06)	0.26 (0.05)	0.31 (0.04)
D. melanogaster, *D. simulans*, *D. mauritania* EST6	0	0	0	0	0	0.09 (0.08)	0.09 (0.11)	0.06 (0.04)	0	0	0.03 (0.02)	0.03 (0.01)	0.04 (0.01)
D. simulans polymorphism EST6	0	0	0	0	0	0.08 (0.08)	0.07 (0.09)	0.05 (0.04)	0	0	0	0	0.02 (0.01)
D. melanogaster polymorphism EST6	0	0	0	0	0	0.07 (0.07)	0	0.03 (0.03)	0.001 (0.03)	0	0	0.01 (0.01)	0.01 (0.01)

tains only six across all the residues that the model predicts would be in the gorge. The three substrate binding regions largely comprise α helices and overall they are no less divergent than other α helices. In absolute terms Table 3 shows substrate binding region 3 to be the most divergent, but the graphical plot in Figure 4 also shows a marked contrast between the divergence in region 1 and the high conservation of the primary sequence around it.

The differences between EST6 and AChE in the gorge and substrate binding regions would be expected to contribute substantially to their differences in substrate utilization. This prediction has been confirmed empirically for selected residues (Gibney *et al.*, 1990; Myers, Healy & Oakeshott, 1993). For example EST6 and AChE have His and Glu respectively at the residue immediately adjacent the Ser nucleophile on its amino terminal side. A synthetic mutant of EST6 with Glu at this position acquires some activity for the diagnostic AChE substrate acetylthiocholine (Myers, Healy & Oakeshott, 1993).

EST6 from D. melanogaster *versus EST5 from* D. pseudoobscura

Although there is good evidence that the two enzymes are orthologous, the comparison of *D. melanogaster* EST6 with EST5 from *D. pseudoobscura* nevertheless entails a qualitative shift in the function of the enzyme. As outlined earlier, EST5 generally exists as a homodimer whereas EST6 in *D. melanogaster* is a monomer. Also EST5 shows little activity in the male reproductive tract but high levels of activity in eyes. Consistent with the implied functional change, there is a high level of replacement site sequence divergence between the two enzymes. The level is 4-5 fold lower than the paralogous EST6/AChE comparison above but, at 23% of the presumptively unconstrained silent site divergence (Brady, Richmond & Oakeshott, 1990), it is still very high compared to the values below 10% obtained for most other orthologous comparisons among *Drosophila* species (Karotam, Delves & Oakeshott, 1993). While some of this replacement variation may be adaptively neutral, it also seems reasonable to conclude that some proportion is responsible for the differences in the function of the enzyme between *D. pseudoobscura* and *D. melanogaster*.

Table 4. The relationship between primary sequence and functional regions as used in Table 3. Nomenclature and identification of functional regions in AChE are taken from Cygler *et al.* (1993). The EST6 nomenclature is based on the model of Cygler *et al.* (1993) and the alignment of EAvP. See also Figures 4 through 7. Gaps or structures not present are indicated by a dash. See Table 1 for identification of beta sheets and alpha helices.

Gorge residues		*Substrate binding*	
AChE	EST6	AChE	EST6
Catalytic triad		Site 1	
Ser-200	Ser-188	71-86	69-77
Glu-327	Asp-319	Site 2	
His-440	His-445	251-264	237-251
Aromatic guidance		Site 3	
Tyr-70	Trp-68	270-278	260-268
Trp-84	-		
Trp-114	His-105	*Bridges*	
Tyr-121	Met-112	Salt Bridges	
Tyr-130	His-121	Arg-44	Arg-41
		Glu-92	Glu-82
Trp-233	Trp-221	Arg-149	Arg-138
		Asp-172	Asp-160
Trp-279	Arg-266		
Phe-288	Pro-275	Glu-163	Asp-151
		Arg-267	Lys-254
Phe-290	Ala-277	Asp-397	Asp-394
Phe-330	Tyr-322	Arg-517	Glu-521
Phe-331	Asn-323	Disulfide linkages	
Tyr-334	Leu-326	Cys-67	Cys-65
		Cys-94	Cys-84
Trp-432	Ala-430	Cys-254	Cys-240
Tyr-442	Asp-447	Cys-265	Cys-252
Other		Cys-402	-
		Cys-521	-
Asn-66	Ala-64	-	Cys-493
Val-71	-	-	Cys-514
Asp-72	Asp-69		
Ser-122	Phe-113		
Glu-199	His-187		
Glu-273	Glu-260		
Leu-282	Leu-269		
Asp-285	Ser-272		
Ser-286	Tyr-273		
Ile-287	Val-274		
Ile-439	Val-444		
Glu-443	Asp-448		
Ile-444	Tyr-449		

This proposition is supported by an examination of the nature and location of the 138 amino acid differences between the *D. melanogaster* and *D. pseudoobscura* enzymes (Brady, Richmond & Oakeshott, 1990). Thus a relatively high proportion of these differences are physicochemically non-conservative (39% for charge and size, and 33% for polarity and hydrophobicity, on the criteria of Taylor, 1986). Likewise, an unusually high proportion of the differences (12%) occur in hydrophobic regions of a hydropathy plot (H > 0.5) which are likely to be in the interior of the protein where a change might have more radical effects on structure. It is also noteworthy that the hydropathy plots for the two enzymes diverge substantially in the 50 residues at their carboxy termini. There is one segment in this region in particular where 8 of 11 contiguous residues involve a charge difference.

Reference to the structural model for carboxyl/cholinesterases in Table 3 and Figure 4 shows that the overall distribution of differences between EST6 and EST5 is similar to that from the previous comparison of EST6 and AChE. Two relatively conserved elements in the EST6/EST5 comparison are the catalytic triad and the salt and cysteine bridges, as they were in the EST6/AChE comparison. Conversely, the active site gorge and substrate binding region are again the most divergent regions. As with the EST6/AChE comparison, this is consistent with a qualitative shift in esterase function between EST6 and EST5. Significantly, however, the nature of the differences between EST6 and EST5 in the gorge residues in particular is not the same as those between EST6 and AChE. For example, while the proportion of aromatic residues in the gorge is much less in EST6 than AChE, it remains about the same for EST5 as compared with EST6. In fact, all aromatic residues are identical between EST6 and EST5 throughout the gorge. This similarity may exemplify an aspect of substrate utilisation which EST6 and EST5 share with each other but not with AChE.

Two other structural features against which we can compare EST6 and EST5 are potential glycosylation sites and the signal peptide. EST6 is known to utilise four glycosylation sites and fur-

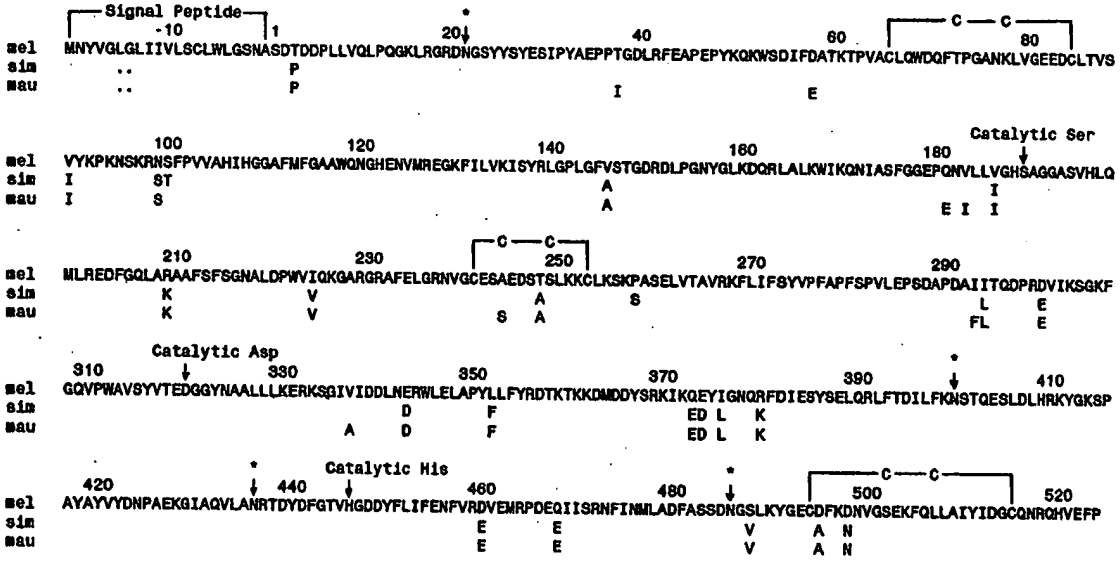


Fig. 5. Inferred EST6 sequences for *D. melanogaster* (mel), with differences in *D. simulans* (sim) and *D. mauritiana* (mau) shown below. Gaps are represented by dots and residues involved in the catalytic triad or glycosylation sites (asterisked) are indicated by arrows. The signal peptide and three cysteine bridges (C-C) are shown above the sequence. Data are from Figure 2 of Karotam, Delves and Oakeshott (1993) but correct some errors in that figure.

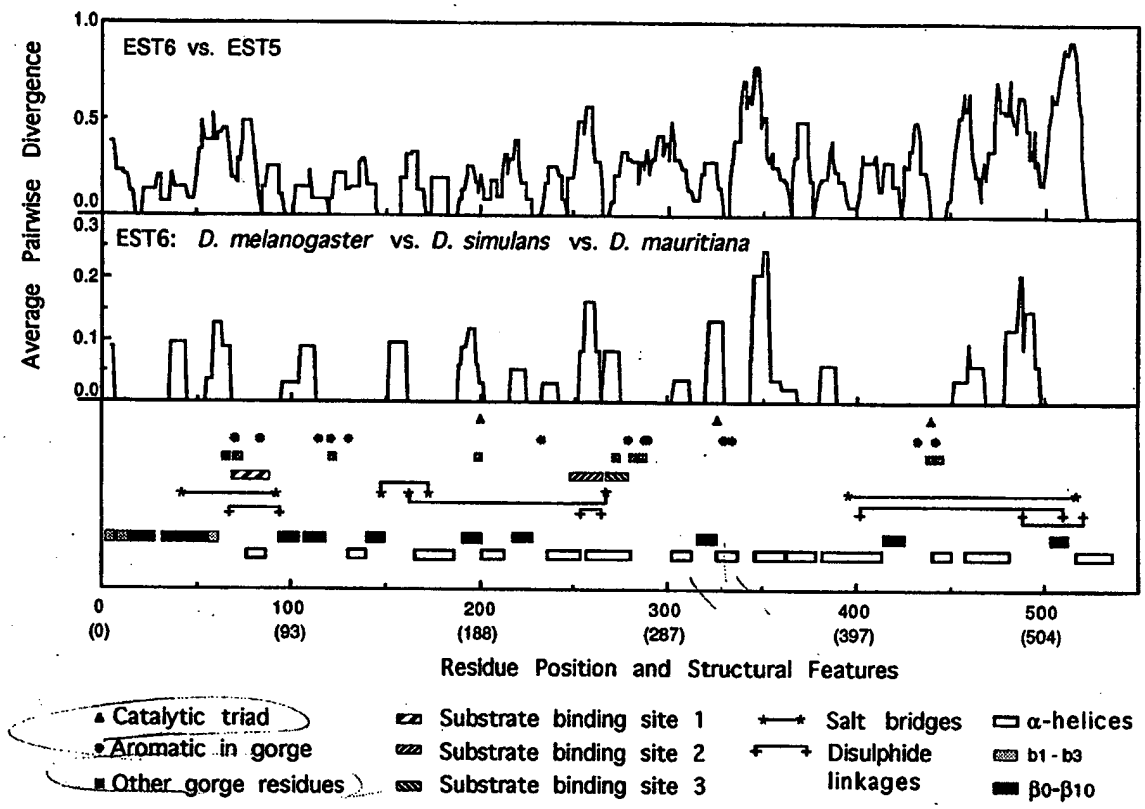


Fig. 6. Average pairwise distances between the mature protein sequences for EST6 from *D. melanogaster*, *D. simulans* and *D. mauritiana*. Distances are calculated and displayed using the same procedures and conventions as Figure 4. The sliding window plot for EST6 from *D. melanogaster* and EST5 from *D. pseudoobscura* from Figure 4 is given for comparison.

thermore, the glycans are known to stabilise the protein *in vivo* (Myers, Healy & Oakeshott, submitted). Glycosylation has not been studied in EST5 but we note that only two of the four glycosylation sites in the *D. melanogaster* protein are conserved in *D. pseudoobscura* and the latter also has one other potential glycosylation site not found in *D. melanogaster*. Over half the 21 residues in the signal peptide differ between the two proteins, although both retain the hydrophobicity characteristics associated with functional secretion signals. Again, the functional significance of secretion is not known for EST5 but it is critical to the reproductive tract function of EST6 (see below).

As we noted earlier, the qualitative differences in structural aspects of the enzyme in *D. melanogaster* and *D. pseudoobscura* are associated with several qualitative differences in its expression. It might therefore be expected that promoter sequences, like the coding region, would also have diverged substantially between the two species. This expectation is fulfilled. The promoter for *Est6* extends about 1.1 kb 5′ of the translation start site (Ludwig, Tamarina & Richmond, 1993) but the *Est5* promoter is barely half this length, extending no further than 450 bp 5′ of the coding region (Brady & Richmond, 1990). Moreover, sequence divergence is so great that the alignment across some segments of the two promoters is no better than chance expectations (Brady, Richmond & Oakeshott, 1990). In general terms the divergence is lower near the start of the coding region, which is consistent with results from a fine-scale dissection of the *D. melanogaster* promoter showing a greater concentration of specific promoter elements in this region (Ludwig, Tamarina & Richmond, 1993). It is also consistent with

the finding that many of the elements required for the *D. melanogaster*-specific ejaculatory duct expression lie in the distal part of that species' promoter. Elements directing the *D. pseudoobscura*-specific expression in the eye have not yet been localised.

Molecular bases for esterase evolution: within function comparisons

EST6 in D. melanogaster *versus* D. simulans *and* D. mauritiana

Unlike the two comparisons above, a comparison of EST6 among the sibling species *D. melanogaster*, *D. simulans* and *D. mauritiana* does not involve any qualitative shift in EST6 function. Extensive comparative data on the biochemistry and physiology of EST6 in these three melanogaster subgroup species reveal some quantitative differences in EST6 activity levels but no qualitative change in its tissue or temporal distribution (Morton & Singh, 1985; Karotam & Oakeshott, 1993, and references therein). The enzyme also takes a monomeric structure in all three species (Morton & Singh, 1985; Oakeshott, Healy & Game, 1990). Consistent with the implied conservation of function, the rate of replacement site divergence between *D. melanogaster* EST6 and either of the other two species' enzymes is lower than the corresponding rate from the comparison between *D. melanogaster* and *D. pseudoobscura* above (14% versus 23% of the respective silent site rate; Karotam, Delves & Oakeshott, 1993). The 30 differences in the mature EST6 proteins of the three melanogaster species are also less radical in physicochemical terms than the differences between the *D. melanogaster* and *D. pseudoobscura* proteins (17% versus 39% non-conservative for charge, 33% versus 39% for size, 23% versus 33% for polarity and hydrophobicity, on the criteria of Taylor, 1986). Moreover, only one of the 30 changes (a conservative Val/Ile difference at residue 182) lies in a hydrophobic sequence ($H > 0.5$), whereas 12% of the differences between the *D. melanogaster* and *D. pseudoobscura* proteins lie in hydrophobic regions.

Further evidence that the amino acid differences in EST6 between the three sibling species may have relatively little effect on function comes from overlaying them on the general model for carboxyl/cholinesterase structure (Table 3, Figs. 5 and 6). Whereas the divergent residues in the cross-functional EST6/AChE and EST6/EST5 comparisons were widely distributed across structural elements, the differences in EST6 among the sibling species are now confined to certain structural features. Most notably, the conservation of the catalytic triad between EST6 and EST5 now extends to all residues identified in the active site gorge and the triad residues are also embedded in invariant stretches of at least ten amino acids in the primary sequence. This suggests tight conservation of the catalytic mechanism across the three species. Other structural features whose residues are invariant in this comparison include the minor β sheet and all the salt and cysteine bridges. The conservation of salt and cysteine bridges suggests conservation of broad structural and stability parameters of the enzyme.

Regions that remain variable among the three species are the substrate binding elements, the major β sheet, the α helices and the regions to which no specific structures are assigned by the model. The divergence in many of these elements may have little impact on function but some of the variability in substrate binding sites 2 and 3 might be expected to affect substrate specificities. Whether this is true awaits empirical data comparing EST6 substrate specificities across the three species.

Two other elements of known function that also vary across the three sibling species involve a glycosylation site and the signal peptide (Karotam, Delves & Oakeshott, 1993). Three of the four glycosylation sites found in *D. melanogaster* recur in the other two species and neither of the latter show any new sites. However the fourth site, residue 487 in *D. melanogaster*, is absent in *D. simulans* and *D. mauritiana*. This difference could affect the *in vivo* stability of the enzyme, although we note that this site is also polymorphic in *D. melanogaster*, so it is not a fixed difference between the species (Myers, Healy & Oakeshott, 1993). The hydrophobic core of the signal peptide is also two residues shorter in *D. simulans* and *D. mauritiana* than in *D. melanogaster* (ten versus twelve residues). This difference could affect the secretion of EST6, although both forms of the signal peptide are within the bounds expected to be functional.

In contrast to the relatively high level of replacement site variation in the *Est6* coding region, the

promoter shows relatively little variation among the three melanogaster subgroup species. Most of the tissue-specific elements in the *D. melanogaster* promoter lie in the first 300-400 bp 5′ of the gene, although some elements controlling male reproductive tract expression lie in the next 600-700 bp (Ludwig, Tamarina & Richmond, 1993, and see above). Preliminary analysis of the *D. simulans* and *D. mauritiana* promoters in transgenic *D. melanogaster* suggests that the qualitative similarities and quantitative differences in EST6 expression between the species are encoded by the first 1.1 kb of 5′ DNA (Karotam, Delves & Oakeshott, 1993). Sequence divergence across the three species is quite limited in the proximal third of this segment (15% of the silent site rate for *D. melanogaster* versus the other two species) but much less constrained in the remainder (75%; Karotam, Delves & Oakeshott, 1993). Features of the proximal segment that are absolutely conserved include the TATA box and two stretches of over 100 bp which contain elements directing expression in four distinct tissues. The conservation of this proximal third is consistent with the qualitative similarity in the temporal and tissue specificity of *Est6* expression across the three species. On the other hand, the divergence in the remainder of the promoter may explain the quantitative differences in EST6 activities among the three species.

Polymorphisms within D. melanogaster *and* D. simulans

Before discussing the molecular bases of polymorphisms in the structure and expression of EST6 in *D. melanogaster* and *D. simulans*, we first summarise the wealth of data on their classical population genetics. The major reason for the interest in these polymorphisms by population geneticists has been the physiological evidence that EST6 contributes to fitness through effects on reproductive behavior. Most of the major pulse of EST6 expression in the anterior sperm ejaculatory duct of the adult male is transferred to the female during mating. Females mating with wild type males show greater oviposition behavior and slower receptivity to remating than do females mating with males homozygous for a laboratory mutant lacking any EST6 activity (Richmond *et al.*, 1990, for a review; see also Miekle & Richmond, 1990; Myers, Healy & Oakeshott, submitted).

Allozyme polymorphism

D. melanogaster is polymorphic for two major allozymes, denoted EST6-F and EST6-S. These variants show complementary latitudinal clines, such that EST6-S frequencies tend to increase at the expense of EST6-F frequencies at higher latitudes. The relationship extends over 40 °C of latitude and is broadly consistent across different continents and both hemispheres (Oakeshott *et al.*, 1981; but see also Jiang, Gibson & Chen, 1989). Consistent with these clines, weak but recurrent seasonal trends have also been recorded, showing EST6-S frequencies to increase at the expense of EST6-F in cooler seasons (Franklin, 1981; Oakeshott, Wilson & Knibb, 1988). *D. simulans* is also polymorphic for EST6-F and EST6-S allozymes and the electrophoretic mobilities of these variants are identical to the corresponding *D. melanogaster* allozymes under standard electrophoretic conditions. Moreover, parallel latitudinal clines are also found for the *D. simulans* variants (Anderson & Oakeshott, 1984). All these data suggest that natural selection differentiates between EST6-F and EST6-S and that the molecular target and mechanism of the selection are shared between the two sibling species.

The first major complication for this interpretation comes from evidence for additional structural variants segregating within EST6-F and EST6-S. This variation is cryptic to the standard electrophoretic procedures used to describe the clines but can be detected by thermostability analyses or higher resolution electrophoresis (Cochrane & Richmond, 1979; Cooke, Richmond & Oakeshott, 1987). In *D. melanogaster* there are several relatively common variants within EST6-F but one form, denoted EST6-8, dominates in frequency within EST-S in all populations investigated (Labate *et al.*, 1989). In *D. simulans* several variants have been reported within both EST6-F and EST6-S but none of them correspond to EST6-8 (Albuquerque & Napp, 1981; Karotam, Boyce & Oakeshott, in press). Two distinct selective processes have therefore been proposed (Oakeshott *et al.*, 1989): one targets the overall EST6-F/EST6-S difference and explains the clines in both species, while the other distinguishes EST6-8 within EST6-S and accounts for its proliferation in *D. melanogaster*.

A second difficulty in further elucidating the selection among EST6 allozymes is the overall failure of about 40 attempts to detect the selection in laboratory populations of *D. melanogaster* (Oakeshott *et al.*, 1989; Richmond *et al.*, 1990, for references). While many studies have reported fitness differences associated with the major EST6-F/EST6-S difference there is little consistency in the effects seen across different studies. Notably few studies have specifically examined reproductive components of fitness that might be relevant to the physiological function of the enzyme in the ejaculatory duct; nevertheless the few that have, have failed (Saad *et al.*, submitted). Importantly also, only two studies have distinguished the minor variants segregating within EST6-F and EST6-S, but neither found any evidence for a fitness advantage to EST6-8 (Saad *et al.*, submitted; Oakeshott *et al.*, submitted). Notwithstanding the methodological problems with some of the studies, the overall conclusion must be that the selection among EST6 allozymes inferred from the field data is either inoperative, or too weak to detect, in the laboratory.

Amino acid polymorphism

Molecular analyses of the various EST6 allozymes can explain some of these complications and anomalies but they also radically change our interpretation of the selection inferred from the field data on allozyme frequencies. Seventeen isolates of the *Est6* gene covering fourteen of the allozymes detectable by high resolution electrophoresis have now been sequenced in *D. melanogaster* and *D. simulans* (Cooke & Oakeshott, 1989; Karotam, Boyce & Oakeshott, in press). The high level of amino acid polymorphism revealed puts EST6 among the most polymorphic isozymes yet characterised at a molecular level in any species. On average, any two of the EST6 allozymes from *D. melanogaster* differ by about four amino acids, and the corresponding figure for *D. simulans* is seven. Several of the amino acid polymorphisms prove to have no mobility phenotype even under high resolution electrophoresis or isoelectric focussing (P.H. Cooke & J.G.O., unpublished). Clearly some of the anomalies in the laboratory fitness comparisons among *Est6* genotypes could simply reflect the use of different EST6 variants.

Surprisingly, no single amino acid polymorphism is invariantly associated with the EST6-F/EST6-S difference across the thirteen sequenced isolates from *D. melanogaster* which have been sequenced. However, most of the EST6-F group of allozymes can be distinguished from most of the EST6-S group by two amino acid polymorphisms in tight linkage disequilibrium with each other, Asp/Asn-237 and Thr/Ala-247 (with the EST6-F amino acid given first). We suggest that the EST6-F/EST6-S clines in *D. melanogaster* are an imperfect reflection of selection on either or both of these amino acid differences. Of these two, the Asp/Asn-237 would be the best candidate for selection, partly because it involves a charge difference that presumably causes the electrophoretic mobility difference, and also because its association with the EST6-F/EST6-S difference is slightly stronger than is that for Thr/Ala-247 (see below).

The data set for *D. simulans* only includes four isolates but this is sufficient to show that EST6-F and EST6-S in this species are not distinguished by the same polymorphisms as those most strongly associated with the electrophoretic difference in *D. melanogaster*. In *D. simulans* the two polymorphisms distinguishing the EST6-F and EST6-S sequences are Thr/Asn-237 and Asp/Val-487, the only overlap with the two in *D. melanogaster* above being the Asn-237 in EST6-S. Therefore, the earlier conclusion of a common target and mechanism of selection in the two species can only be retained by proposing that selection would recognise Asp-237 and Thr-237 in the two species' EST6-Fs as differing in the same way from Asn-237 in the two EST6-Ss. This does not seem likely for two reasons. Firstly, it seems unlikely on physicochemical grounds, since Asp is charged but Thr not. Secondly, since only the Asp/Val-487 of the two *D. simulans* polymorphisms involves a charge difference, only this 487 polymorphism is likely to be causally connected with the electrophoretic mobility difference. Given the distance between the two polymorphisms and the fact that only four *D. simulans* sequences were sampled, the association of Thr/Asn-237 with the EST6-F/EST6-S may not even hold up in the wider population.

Another revision to our thinking caused by the molecular data concerns the EST6-8 allozyme that is common in *D. melanogaster* but absent from *D. simulans*. On the assumption that the EST6-F/EST6-S difference has the same molecular basis in the two species, the proliferation of EST6-8 within

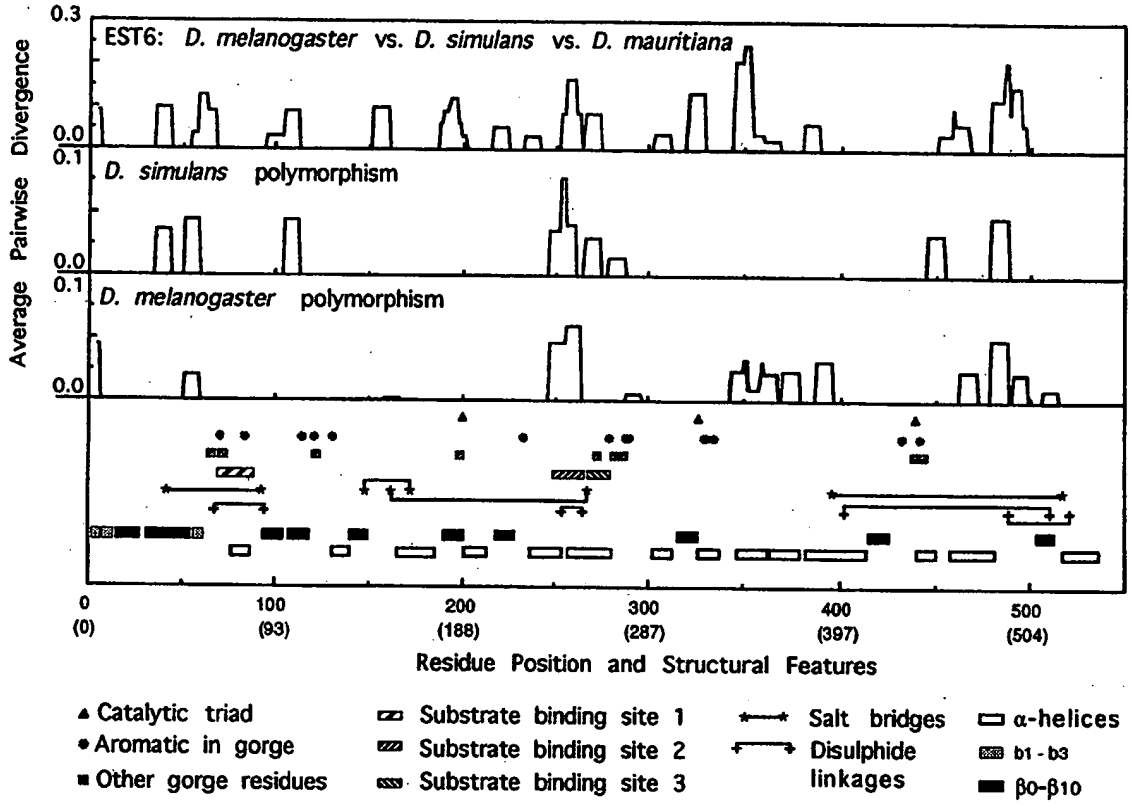


Fig. 7. Average pairwise distances between the mature protein sequences of 12 *Est6* alleles from *D. melanogaster* and four from *D. simulans*. Distances are calculated as the average over all alleles within each species. Conventions otherwise follow Figures 4 and 6, with the sliding window plots for divergence among *D. melanogaster*, *D. simulans* and *D. mauritiana* given for comparison.

EST6-S in *D. melanogaster* was originally attributed to the spread of a different amino acid polymorphism from those underlying the clinal selection between EST6-F and EST6-S. Since the assumption about EST6-F/EST6-S is almost certainly wrong, there is no need to propose that the demographies of the EST6-8 and EST6-S variants in *D. melanogaster* reflect the demographies of different amino acid polymorphisms. Indeed, given that EST6-8 is by far the most common allozyme within the EST6-S group, they probably reflect effects on the same amino acid polymorphisms.

One of the original ideas about EST6-8 that is supported by the molecular data is the notion that it has arisen and spread relatively recently. Sequencing of two *Est6-8* isolates from different continents reveals no replacement or silent site differences between them, whereas replicate isolates for two other allozymes differ at an average of seven such sites. Importantly however, the molecular data also show that *Est6-8* is just a representative of a small group of four haplotypes that are closely related to one another (average of three differences in replacement and silent sites) but distantly related to the eight other haplotypes detected (average of twelve differences). The EST6-8 group contains three EST6-S haplotypes and a rare faster allozyme (EST6-F′), while the other group includes all the EST6-F forms, another rare very fast mobility form (EST6-vF), and the fastest mobility form within EST6-S. The only polymorphism invariably associated with the two groups of haplotypes is Asp/Asn-

237 but Thr/Ala-247 and several silent polymorphisms less than 300 bp away are also strongly associated with the two haplotype groupings (the only exception for 247 being Ala-247 in one of the EST6-8 group of haplotypes). While the 237 polymorphism thus remains the most likely target for the selection inferred from the EST6-F/EST6-S clines and the proliferation of EST6-8, the response to selection has nevertheless involved a block of several closely linked polymorphisms. It is at least a formal possibility that some of the silent polymorphisms could influence phenotype and fitness via effects on the regulation of the *EstP* gene immediately downstream.

Functional effects of the amino acid polymorphisms

How then do the various amino acid polymorphisms in *D. melanogaster* and *D. simulans* relate to the structural model for carboxyl/cholinesterases? As was the case with the interspecific comparisons among *D. melanogaster*, *D. simulans* and *D. mauritiana* (Fig. 7, Table 3), the active site gorge, salt and cysteine bridges and the minor β sheet are essentially invariant (see below for one exception). Thus, these structures are divergent in comparisons among enzymes with qualitatively different functions (i.e., EST6/AChE and EST6/EST5) but conserved in comparisons of enzymes with qualitatively similar functions (i.e., EST6 within and among the sibling species). On the other hand, there is also relatively greater conservation in the major β sheet and many of the α helices than was evident in the comparisons between the sibling species. The vast majority of the polymorphism lies either in substrate binding sites or in regions to which no particular structures are ascribed in the model.

In fact, only seven out of the total of 32 polymorphisms lie in structures to which specific functional effects can be attributed, four in substrate binding site 2, two in a glycosylation site and one in a cysteine bridge. The latter, a Cys/Tyr-514 polymorphism in *D. melanogaster*, is the cause of an extremely rare, smeary electrophoretic phenotype that was specifically chosen for sequencing because of this phenotype but which we have only ever noticed in a single field strain. The other six of these polymorphisms are all relatively common. There are three of them in each species and, significantly, they include both polymorphisms most strongly associated with EST6-F and EST6-S in each species. Because two of the sites for these six polymorphisms are shared across the two species, the six differences only involve four sites, 237, 243, 247 and 487.

The Asp/Asn-237 and Thr/Ala-247 polymorphisms which are most strongly associated with the EST6-F/EST6-S difference in *D. melanogaster* both lie in substrate binding site 2. Asp/Asn-237 produces the charge difference that presumably causes the electrophoretic mobility difference while Thr/Ala-247 is also non-conservative for polarity. Some effects on substrate interactions would therefore seem possible for both polymorphisms, although they are arguably more likely for the charge-non-conservative Asp/Asn-237. Differences in kinetic parameters related to substrate interactions and in substrate and (competitive) inhibitor specificities have indeed been reported between purified EST6-F and EST6-S (Danford & Beardmore, 1979; White, Mane & Richmond, 1988). This further supports the proposition that the 237 and/or 247 polymorphisms are targets for the natural selection that underlies the clines for EST6-F/EST6-S (and the rapid proliferation of EST6-8). The third common polymorphism in *D. melanogaster* that the structural model suggests might affect function is Ser/Ala-487, which results in the presence/absence of the fourth and final glycosylation site in the EST6 primary sequence. As explained earlier, EST6 mutants engineered to lack this plus its three other glycosylation sites are less stable *in vivo*. Consistent with this, the Ser/Ala-487 difference is associated with minor electrophoretic and thermostability variation within both EST6-F and EST6-S (Cooke & Oakeshott, 1989).

As in *D. melanogaster*, two of the three common *D. simulans* polymorphisms adjudged on structural grounds most likely to affect function lie in substrate binding site 2 and the third lies in the fourth glycosylation site above. In *D. simulans* the binding site polymorphisms are Thr/Asn-237 and Ala/Ser-243, the first of which is associated with the EST6-F/EST6-S difference. Thr/Asn-237 is non-conservative for hydrophobicity while Ala/Ser-243 is non-conservative for polarity and hydrophobicity. Effects on substrate interactions therefore seem possible for both, although we are unaware of any empirical data to test this possibility. The third

D. simulans polymorphism in this group is Val/Asp-487, which produces the charge difference assumed to underlie the EST6-F/EST6-S electrophoretic mobility difference in this species. Residue 487 aligns with the glycosylation site in *D. melanogaster* EST6 but neither amino acid at this position in the *D. simulans* protein would in fact produce a glycosylation site. However, Val/Asp-487 is non-conservative for size and hydrophobicity as well as for charge. That such a physicochemically radical difference occurs at a residue where glycosylation affects the function of the *D. melanogaster* protein suggests the possibility of functional effects in *D. simulans* EST6, albeit ones we cannot yet specify.

Promoter polymorphism

Although the analyses are less advanced than for the structural region, early data suggest that regulatory polymorphisms also contribute substantially to EST6 phenotypic variation within *D. melanogaster* and *D. simulans*. Heritable 2-3 fold differences in the V_{max} of EST6 have been reported among 42 third chromosome isoallelic lines extracted from a natural population of *D. melanogaster* (Game & Oakeshott, 1989). Differences of a similar order have also been recorded in seventeen of these lines scored for EST6 V_{max} in first instar larvae and mid-pupae of this species (Oakeshott *et al.*, submitted). Only a small minority of the variation in any of the measures can be explained by differences among the six EST6 allozymes that high resolution electrophoresis shows to segregate among the lines. More of the activity variation is associated with restriction fragment length polymorphisms (RFLPs) located 5′ of the gene.

More specifically, all significant between-line variation in larval and pupal activities is associated with polymorphism for the insertion of a transposable element about 1.4 kb 5′ of the gene (Oakeshott *et al.*, submitted). We have seen that most of the Est6 promoter lies closer to the gene, so this effect may be an indirect one, due to interference in the promoter's functions during transcription of the transposable element. No 5′ RFLPs have been related to the variation in adult female activity levels but about 20% of the differences between lines in male activity levels is associated with a *Rsa*I polymorphism about 540 bp 5′ of the coding region (Game & Oakeshott, 1990). Promoter elements directing *Est6* expression in adult males are known to lie in or near the *Rsa*I site. However, without full sequence data it is not possible to determine whether the *Rsa*I polymorphism is causally related to the activity variation, or simply in linkage disequilibrium with other promoter polymorphism(s) which are more directly involved.

Although only seven *D. simulans* lines have been compared for EST6 V_{max}, high levels of variation among the lines are evident in both adult males and females (Karotam & Oakeshott, 1993). The *Est6* promoter has also been sequenced for four of these lines (Karotam, Boyce & Oakeshott, in press). Although the data set is too small to relate the activity differences to particular promoter polymorphisms, some broad patterns are manifest in the distribution of the nucleotide variation along the promoter. Polymorphism is extremely low in the proximal third of the 1.1 kb of 5′ DNA sequenced (6% of the silent site level in the coding region), but much higher in the remaining two thirds (41% of the silent site level). This distribution closely parallels the pattern for the same region we saw earlier in the interspecific comparisons among *D. melanogaster*, *D. simulans* and *D. mauritiana*. However, as with these earlier comparisons, the localisation of the particular sequence differences causing the activity variation will await more detailed dissection of *Est6* promoter function.

Conclusions

The aim of this essay has been to reconstruct some of the macro- and micro-evolutionary processes shaping the diversity of esterase enzymes found in higher organisms. The focus of our attention is *Drosophila*, *D. melanogaster* in particular, because of the wealth of data on its biochemical, molecular and evolutionary genetics. However, our interpretation of these data relies heavily on functional classification schemes and structure-function models originally developed with vertebrate and microbial esterases.

About 30 distinct esterases have been detected in *D. melanogaster* by electrophoretic analyses and several others have been identified from this species by spectrophotometric methods. Most can be classified as carboxyl, choline or acetyl esterases on a set of inhibitor-based criteria originally devel-

oped for mammalian esterases. A fourth class found in mammals, the aryl esterases, has not yet been found in *D. melanogaster*, although it is not clear that the assays used with *D. melanogaster* have been appropriate to detect such a class. The preponderance of carboxyl and cholinesterases in both mammals and *D. melanogaster* contrasts with the finding that most prokaryotic esterases are aryl or acetyl esterases or, more often, not classifiable against the inhibitor criteria. Thus, esterases can be partitioned into several families on *in vitro* biochemical criteria but there is only limited overlap between the families that have proliferated in eukaryotes versus prokaryotes.

Sequence data for over fifty cloned esterase genes establish that at least some of these biochemically defined families represent distinct phylogenetic lineages as well. There are significant similarities between many of the eukaryotic carboxyl esterases and all the eukaryotic cholinesterases sequenced to date, so these enzymes are all placed in a single carboxyl/cholinesterase multigene family. This family encompasses all the *Drosophila* esterases so far sequenced. There is minimal sequence similarity between the carboxyl/cholinesterase and any other esterases but crystal structures for selected enzymes reveal striking similarities with some other mammalian carboxyl esterases, mainly lipases, and some bacterial aryl esterases. This suggests that a high proportion of eukaryotic esterases and at least some prokaryotic esterases share a common, albeit ancient, ancestry. All the enzymes that show the characteristic tertiary structure, or bear sequence similarity to enzymes with this structure, are placed in a superfamily, called the α/β hydrolases. The affinities of the remaining enzymes, mostly aryl or acetyl esterases or unclassifiable on the inhibitor criteria, are generally unknown, although a few of the unclassifiable enzymes show sequence similarity with serine proteases.

About fifteen esterase genes have been mapped by classical or molecular methods to the *D. melanogaster* genome. Five separate chromosomal sites are implicated, with two genes so far identified at one site, and about ten at another. The latter cluster, termed the α-cluster, includes both carboxyl and cholinesterases, so its origin may pre-date the carboxyl/cholinesterase split. Certain members of this family are tightly conserved across diverse *Drosophila* species and preliminary molecular work also suggests orthology with esterases implicated in insecticide resistance in several other insects. By contrast, the other, smaller cluster, seems to be evolving more rapidly. Both genes in this cluster, termed the β-cluster, are found in other members of the subgenus Sophophora, although the expression of one gene varies qualitatively in *D. pseudoobscura* and that species' β-cluster also contains a third esterase sequence. Evidence to date suggests that the number, organisation and expression of genes in the β-cluster all vary in more distantly related species from the subgenus *Drosophila* and there is indirect evidence for two β-clusters in some of those species.

We have used the EST6 carboxyl esterase from the β-cluster to investigate the molecular basis for functional evolution within the carboxyl/cholinesterase multigene family. Preliminary modelling shows good agreement with the super-secondary structure for a vertebrate cholinesterase, AChE, whose crystal structure has been solved. However, primary sequence differences between the two enzymes are widespread throughout the super-secondary structure. Divergence is lowest around salt and cysteine bridges that stabilise the overall α/β hydrolase fold structure. On the other hand, individual residues in the α helices and β strands that make up this overall structure are quite variable. Indeed these structures are no less variable than regions largely expected to be on the protein surface to which no particular structures are ascribed by the model. High levels of divergence are also evident in substrate binding regions and the active site gorge and these may be critical to the functional differences between the two enzymes. In particular, most of the aromatic guidance residues in the active site gorge of AChE do not occur in EST6.

Comparisons of EST6 from *D. melanogaster* with the orthologous EST5 from *D. pseudoobscura* also involve a qualitative change in function, the former being mainly expressed in male reproductive tissue and the latter in the eye. This difference in function is again associated with a high level of sequence divergence; although the two species lie in the same subgenus, the two enzymes differ at about a quarter of their residues. As with the EST6/AChE comparisons, the sequence differences are also widespread across the predicted super-secondary structure. Salt and cysteine bridges are again

conserved but high levels of divergence occur in the α helices, the β strands in both β sheets, the substrate binding sites, the active site gorge and regions for which no particular higher order structure is identified.

Unlike the EST6/AChE and EST6/EST5 comparisons, comparisons of EST6 between *D. melanogaster* and its two sibling species *D. simulans* and *D. mauritiana* involve a high level of functional conservation. Although sequence divergence is again high compared to other enzymes, the locations of the differences in the super-secondary structure are now more restricted. The relatively conserved elements now include the active site gorge and the smaller of the two β sheets, as well as the salt and cysteine bridges. Variation is concentrated in α helices, the major β sheet, regions with no identified higher order structure and, perhaps surprisingly, the substrate binding sites.

As in the interspecific comparisons, levels of EST6 polymorphism within *D. melanogaster* and *D. simulans* are very high compared to other enzymes. However, the location of the variation is even more restricted than in the comparisons between the sibling species. Variation is now essentially confined to the substrate binding sites and the regions for which no particular higher order structure is identified. It is possible that many of the polymorphisms in the latter regions may be neutral to function and fitness. However, kinetic studies on purified preparations of *D. melanogaster* allozymes differing in their substrate binding sites reveal differences in substrate interactions and specificities. Furthermore demographic data on these allozymes reveal non-random spatial and temporal distributions indicative of natural selection.

Analyses of the respective promoters are less advanced than the structural comparisons of EST6 within and among the various species above. Nevertheless some clear parallels in the pattern of conservation and change emerge. Thus the promoters of *Est6* in *D. melanogaster* and *D. pseudoobscura* show limited sequence similarity and some fundamental differences in organisation. High levels of sequence variation are also evident in comparisons of the *Est6* promoters in *D. melanogaster*, *D. simulans* and *D. mauritiana* but their overall organisation is conserved and much of the first 300 bp of 5′ sequence is absolutely invariant. The same holds true of the promoter polymorphisms within *D. melanogaster* and *D. simulans*. As with the structural polymorphisms, some of the promoter polymorphisms may be irrelevant to function and fitness but particular polymorphisms lie in elements which early functional data indicate are required for specific pulses of expression and demographic data suggest are under selection.

Acknowledgements

We thank David Ollis for instructing us in the details of the α/β hydrolase fold and Phil Batterham and Jill Karotam for their critical comments and advice on various drafts of this review.

References

Abdel-Aal, Y. A. I., T. N. Hanzlik, B. D. Hammock, L. G. Harshman & G. Prestwich, 1988. Juvenile hormone esterases in two heliothines: kinetic, biochemical and immunogenic characterisation. Comp. Biochem. Physiol. 90B: 117-124.

Albuquerque, C. M. R. & M. Napp, 1981. Genetic variability at the esterase 6 locus in natural populations of *Drosophila simulans* in relation to environmental heterogeneity. Genetics 98: 399-407.

Anderson, P. R. & J. G. Oakeshott, 1984. Parallel geographic patterns of allozyme variation in two sibling *Drosophila* species. Nature 308: 729-731.

Arnason, E. & G. K. Chambers, 1987. Macromolecular interaction and the electrophoretic mobility of esterase 5 from *Drosophila pseudoobscura*. Biochem. Genet. 25: 287-301.

Arpagaus, M., D. Fournier & J. P. Toutant, 1988. Analysis of acetyl cholinesterase molecular forms during the development of *Drosophila melanogaster*. Evidence for the existence of an amphiphilic monomer. Insect Biochem. 18: 539-549.

Baker, W. K., 1980. Evolution of the alpha-esterase duplication within the montana subphylad of the virilis species group of *Drosophila*. Genetics 94: 733-748.

Beckman, L. & F. M. Johnson, 1974. Esterase variations in *Drosophila melanogaster*. Hereditas 51: 212-220.

Berger, E. & R. Canter, 1973. The esterases of *Drosophila*. I. The anodal esterases and their possible role in eclosion. Devel. Biol. 33: 48-55.

Brady, J. P. & R. C. Richmond, 1990. Molecular analysis of evolutionary changes in the expression of *Drosophila* esterases. Proc. Natl. Acad. Sci. USA 87: 8217-8221.

Brady, J. P. & R. C. Richmond, 1992. An evolutionary model for the duplication and divergence of esterase genes in *Drosophila*. J. Mol. Evol. 34: 506-521.

Brady, J. P., R. C. Richmond & J. G. Oakeshott, 1990. Cloning of the esterase 5 locus from *Drosophila pseudoobscura* and comparison with its homologue in *Drosophila melanogaster*. Mol. Biol. Evol. 7: 525-546.

Campbell, P. M., M. J. Healy & J. G. Oakeshott, 1992. Char-

acterisation of juvenile hormone esterase in *Drosophila melanogaster*. Insect Biochem. Molec. Biol. 22: 665-677.

Cavener, D., D. C. Otteson & T. C. Kaufman, 1986. A rehabilitation of the genetic map of the 84B-D region of *Drosophila melanogaster*. Genetics 114: 111-123.

Choi, K. D., G. H. Jeohn, J. S. Rhee & O. J. Yoo, 1990. Cloning and nucleotide sequence of an esterase gene from *Pseudomonas fluorescens* and expression of the gene in *Eschericia coli*. Agric. Biol. Chem. 54: 2039-2045.

Chothia, C. & A. M. Lesk, 1986. The relationship between the divergence of sequence and structure in proteins. EMBO J. 5: 823-826.

Coates, P. M., M. A. Mestriner & D. A. Hopkinson, 1975. A preliminary genetic interpretation of the esterase isozymes of human tissues. Ann. Hum. Genet. 39: 1-20.

Cochrane, B. J. & R. C. Richmond, 1979. Studies of esterase-6 in *Drosophila melanogaster*. II. The genetics and frequency distributions of naturally occuring variants studied by electrophonesis and heat stability criteria. Genetics 93: 461-478.

Collet, C., K. M. Nielsen, R. J. Russell, M. Karl, J. G. Oakeshott & R. C. Richmond, 1990. Molecular analysis of duplicated esterase genes in *Drosophila melanogaster*. Molec. Biol. Evol. 7: 9-28.

Cooke, P. H. & J. G. Oakeshott, 1989. Amino acid polymorphisms for Esterase 6 in *Drosophila melanogaster*. Proc. Natl. Acad. Sci. USA 86: 1426-1430.

Cooke, P. H., R. C. Richmond & J. G. Oakeshott, 1987. High resolution electrophoretic variation at the Esterase-6 locus in a natural population of *Drosophila melanogaster*. Heredity 59: 259-264.

Cygler, M., J. D. Schrag & F. Ergan, 1992. Advances in structural understanding of lipases. Biotech. Genet. Eng. Rev. 10: 143-184.

Cygler, M., J. D. Schrag, J. C. Sussman, M. Harel, I. Silman, M. K. Gentry & B. P. Doctor, 1993. Relationship between sequence conservation and three dimensional structure in a large family of esterases, lipases and related proteins. Protein Sci. 2: 366-382.

Danford, N. D. & J. A. Beardmore, 1979. Biochemical properties of esterase 6 in *Drosophila melanogaster*. Biochem. Genet. 17: 1-22.

Dayhoff, M. O., R. M. Schwartz & B. C. Orcutt, 1978. A model of evolutionary change in proteins, pp. 345-352 in Atlas of Protein Sequence and Structure, edited by M. O. Dayhoff. The National Biomedical Research Foundation, Silver Spring, Maryland.

De la Escalera, S., E.-O. Bockamp, F. Moya, M. Piovant & F. Jiménez, 1990. Characterisation and gene cloning of neurotactin, a *Drosophila* transmembrane protein related to cholinesterases. EMBO J. 9: 3593-3601.

Devonshire, A. L., 1977. The properties of a carboxylesterase from the peach-potato aphid, *Myzus persicae* (Sulz.), and its role in conferring insecticide resistance. Biochem. J. 167: 675-683.

Di Persio, L. P., R. N. Fontaine & D. Y. Hui, 1990. Identification of the active site serine in pancreatic cholesterol esterase by chemical modification and site-specific mutagenesis. J. Biol. Chem. 265: 16801-16806.

Di Persio, L. P., R. N. Fontaine & D. Y. Hui, 1991. Site-specific mutagenesis of an essential histidine residue in pancreatic cholesterol esterase. J. Biol. Chem. 266: 4033-4036.

Di Persio, L. P. & D. Y. Hui, 1993. Aspartic acid 320 is required for optimal activity of rat pancreatic cholesterol esterase. J. Biol. Chem. 268: 300-304.

East, P. D., 1982. Non-specific esterases of *Drosophila buzzatii*, pp. 323-338, in Ecological Genetics and Evolution. The Cactus-Yeast-Drosophila Model System, edited by J. S. F. Barker and W. T. Starmer. Academic Press, New York.

East, P. D., 1984. Biochemical Genetics of Two Highly Polymorphic Esterases in *Drosophila Buzzatii*, 211 pp. PhD Thesis. University of New England, Armidale.

East, P. D., A. Graham & G. Whitington, 1990. Molecular isolation and preliminary characterisation of a duplicated esterase locus in *Drosophila buzzatii*, pp. 389-406 in Ecological and Evolutionary Genetics of *Drosophila*, edited by J. S. F. Barker, W. T. Starmer and R. J. MacIntyre. Plenum Press, New York.

Ecobichon, D. J. & W. Kalow, 1965. Properties and classification of the soluble esterases of human skeletal and smooth muscle. Canad. J. Biochem. 43: 73-79.

Enikolopov, G. N., O. A. Malevanchuk, N. I. Peunova, P. V. Sergeev, & G. P. Georgiev, 1989. Est locus in *Drosophila virilis* contains two related genes. Doklady Akad. Nauk. SSSR 306: 1247-1249.

Farmer, J. L. & M. W. Carter, 1989. Properties of partially purified allozymes of esterase 5 of *Drosophila pseudoobscura*. Comp. Biochem. Physiol. 93B: 451-458.

Farrell, D. H., P. Mikesell, L. A. Actis & J. H. Crosa, 1990. A regulatory gene, angR, of the iron uptake system of *Vibrio anguillarum*: similarity with phage P22 cro and regulation by iron. Gene 86: 45-51.

Felsenstein, J., 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39: 783-791.

Felsenstein, J., 1991. Phylip. University of Washington, Seattle, Washington.

Field, L. M., M. S. Williamson, G. D. Moores & A. L. Devonshire, 1993. Cloning and analysis of the esterase genes conferring insecticide resistance in the peach-potato aphid *Myzus persicae* (Sulzer). Biochem. J. (in press).

Fournier, D., J.-B. Bergé, M.-L. C. de Almeida & C. Bordier, 1988. Acetylcholinesterases from *Musca domestica* and *Drosophila melanogaster* brain are linked to membranes by a glycophospholipid anchor sensitive to an endogenous phospholipase. J. Neurochem. 50: 1158-1163.

Franklin, I. R., 1981. An analysis of temporal variation at isozyme loci in *Drosophila melanogaster*, pp. 217-236 in Genetic Studies of *Drosophila* Populations, edited by J.B. Gibson and J.G. Oakeshott. The Australian National University Press, Canberra.

Game, A. Y. & J. G. Oakeshott, 1989. Variation in the amount and activity of Esterase 6 in a natural population of *Drosophila melanogaster*. Heredity 62: 27-34.

Game, A. Y. & J. G. Oakeshott, 1990. Associations between restriction site polymorphism and enzyme activity variation for esterase 6 in *Drosophila melanogaster*. Genetics 126: 1021-1032.

Gan, K. N., A. Smolen, H. W. Eckerson & B. N. la Du, 1991. Purification of human serum paraoxonase/arylesterase: evidence for one esterase catalysing both activities. Drug Metabolism Disposition 19: 100-106.

Gibney, G., S. Camp, M. Dionne, K. Macphee-Quigley & P. Taylor, 1990. Mutagenesis of essential functional residues in acetylcholinesterase. Proc. Natl. Acad. Sci. USA 87: 7546-7550.

Healy, M. J., M. M. Dumancic & J.G. Oakeshott, 1991. Biochemical and physiological studies of soluble esterases from *Drosophila melanogaster*. Biochem. Genet. 29: 365-388.

Heymann, E., 1980. Carboxylesterases and amidases, pp. 291-323 in Enzymatic Basis of Detoxication. Vol II, edited by W.B. Jakoby. Academic Press, New York.

Holmes, R. S. & C. J. Masters, 1967. The developmental multiplicity and isoenzyme status of cavian esterases. Biochim. Biophys. Acta 132: 379-399.

Hjoring, N. & O. Svensmark, 1988. Molecular and catalytic properties of a butyrylesterase from human red cells and brain. Arch. Biochem. Biophys. 260: 351-358.

Hui, D. Y., K. Hayakawa & J. Oizumi, 1993. Lipoamidase activity in normal and mutagenized pancreatic cholesterol esterase (bile salt stimulated lipase). Biochem J. 291: 65-69.

Jiang, C., J. B. Gibson & H. Chen, 1989. Genetic differentiation in populations of *Drosophila melanogaster* from the People's Republic of China: comparison with patterns on other continents. Heredity 62: 193-198.

Johnson, F. M., C. G. Kanapi, R. H. Richardson, M. R. Wheeler & W. S. Stone, 1966. An operational classification of *Drosophila* esterases for species comparison. Univ. Texas Publ. 6615: 517-532.

Johnson, F. M., R. H. Richardson & M. P. Kambysellis, 1968. Isozyme variability in species of the genus *Drosophila*. III. Qualitative comparison of the esterases of *D. aldrichi* and *D. mulleri*. Biochem. Genet. 1: 239-247.

Kambysellis, M. P., F. M. Johnson & R. H. Richardson, 1968. Isozyme variability in species of the genus *Drosophila*. IV. Distribution of the esterases in the body tissues of *D. aldrichi* and *D. mulleri* adults. Biochem. Genet. 1: 249-265.

Kao, L. R., N. Motoyama & W. C. Dauterman, 1985. The purification and characterisation of esterases from insecticide-resistant and susceptible house flies. Pestic. Biochem. Physiol. 23: 228-239.

Karotam, J. & J. G. Oakeshott, 1993. Regulatory aspects of esterase 6 activity variation in sibling *Drosophila* species. Heredity 71: 41-50.

Karotam, J., A. C. Delves & J. G. Oakeshott, 1993. Conservation and change in structural and 5' flanking sequences of esterase 6 in sibling *Drosophila* species. Genetica 88: 11-28.

Karotam, J., T. M. Boyce & J. G. Oakeshott, 1994. Nucleotide variation at the hypervariable esterase 6 isozyme locus of *Drosophila simulans*. Mol. Biol. Evol. (in press).

Korochkin, L. I., M. Z. Ludwig, E. V. Poliakova & M. R. Philinova, 1987. Some molecular genetic aspects of cellular differentiation in *Drosophila*. Sov. Sci. Rev. F. Physiol. Gen. Biol. 1: 411-466.

Korochkin, L., M. Ludwig, N. Tamarina, I. Uspensky, G. Yenikolopov, R. Khechumijan, M. Kopantseva, M. Evgeniev, B. Kuzin, T. Bakayeva, L. Mudjoian, O. Malevantschuk, V. Tsatyian, A. Ivanov & S. Lukianov, 1990. Molecular genetic mechanisms of tissue specific esterase isozymes and protein expression in *Drosophila*, pp. 399-440 in Isozymes: Structure, Function and Use in Biology and Medicine, edited by C. Markert and J. Scandalios. Wiley-Liss, New York.

Korochkin, L. I., N. M. Matveeva & A. Yu. Kerkis, 1973. Subcellular localisation of esterases in *D. virilis* and *D. texana*. Dros. Inform. Serv. 50: 130-131.

Labate, J., A. Bortoli, A. Y. Game, P. H. Cooke & J. G. Oakeshott, 1989. Further observations on the number and distribution of esterase 6 alleles in populations of *Drosophila melanogaster*. Heredity 63: 203-208.

Lai-Fook, J. & P. H. Smith, 1991. Genetics of the hemolymph esterases of *Lucilia cuprina* (*Diptera*: *Calliphoridae*). Biochem. Genet. 30: 123-130.

Loukas, M., 1981. A new esterase locus in *D. melanogaster*. Dros. Inform. Serv. 56: 85.

Ludwig, M. Z., N. A. Tamarina & R. C. Richmond, 1993. Localisation of sequences controlling the spatial, temporal and sex-specific expression of the esterase 6 locus in *Drosophila melanogaster* adults. Proc. Natl. Acad. Sci. USA 90: 6233-6237.

Luthi, E., D. R. Love, J. McAnulty, C. Wallace, P. A. Caughey, D. Saul & P. C. Bergquist, 1990. Cloning, sequence analysis and expression of genes encoding xylan-degrading enzymes from the thermophile *Caldocellum saccharolyticum*. Appl. Env. Microbiol. 56: 1017-1024.

Mane, S. D., C. S. Tepper & R. C. Richmond, 1983. Studies of esterase 6 in *Drosophila melanogaster*. XIII. Purification and characterisation of the two major isozymes. Biochem. Genet. 21: 1019-1040.

Markovic, O. & H. Jornvall, 1986. Pectinesterase: the primary structure of the tomato enzyme. Eur. J. Biochem. 158: 455-462.

Mercken, L., M.-J. Simmons, S. Swillens, M. Massaer & G. Vassart, 1985. Primary structure of bovine thyroglobulin deduced from the sequence of its 8431-base complementary DNA. Nature 316: 647-651.

Miekle, D. B. & R. C. Richmond, 1990. Localisation and longevity of seminal fluid esterase 6 in mated female *Drosophila melanogaster*. J. Insect Physiol. 36: 93-101.

Morton, R. A. & R. S. Singh, 1985. Biochemical properties, homology, and genetic variation of *Drosophila* non-specific esterases. Biochem. Genet. 23: 959-973.

Mouchés, C., M. Magnin, J.-P. Bergé, M. de Silvestri, V. Beyssat, N. Pasteur & G. P. Georghiou, 1987. Overproduction of detoxifying esterases in organophosphate-resistant *Culex* mosquitoes and their presence in other insects. Proc. Natl. Acad. Sci. USA 84: 2113-2116.

Mulley, J. C., J. W. James & J. S. F. Barker, 1979. Allozyme genotype-environment relationships in natural populations of *Drosophila buzzatii*. Biochem. Genet. 17: 105-126.

Myers, M. A., M. J. Healy & J. G. Oakeshott, 1993. Effects of the residue adjacent to the reactive serine on the substrate interactions of Drosophila esterase 6. Biochem. Genet. 31: 259-278.

Myers, M. A., M. J. Healy & J. G. Oakeshott, 1994. Effects of four N-linked glycans on the *in vivo* function of esterase 6 in *D. melanogaster* (submitted for publication).

Narise, S., 1973. Esterase isozymes of *Drosophila virilis*: purification and properties of α- and β-esterase isozymes. Japan. J. Genet. 48: 119-132.

Narise, S. & J. L. Hubby, 1966. Purification of esterase 5 from *Drosophila pseudoobscura*. Biochem. Biophys. Acta. 122: 281-288.

Nei, M., 1987. Molecular Evolutionary Genetics. Columbia University Press, New York.

Oakeshott, J. G., G. K. Chambers, J. B. Gibson & D. A. Willcocks, 1981. Latitudinal relationships of esterase-6 and phosphoglucomutase gene frequencies in *Drosophila melanogaster*. Heredity 47: 385-396.

Oakeshott, J. G., C. Collet, R. Phillis, K. M. Nielsen, R. J. Russell, G. K. Chambers, V. Ross & R. C. Richmond, 1987. Molecular cloning and characterisation of Esterase 6, a serine hydrolase from *Drosophila*. Proc. Natl. Acad. Sci. USA 84: 3359-3363.

Oakeshott, J. G., P. H. Cooke, R. C. Richmond, A. Bortoli, A. Y. Game & J. Labate, 1989. Molecular population genetics of structural variants of esterase 6 in *Drosophila melanogaster*. Genome 31: 788-797.

Oakeshott, J. G., M. J. Healy & A. Y. Game, 1990. Regulatory evolution of β-carboxyl esterases in *Drosophila*, pp. 359-387 in Ecological and Evolutionary Genetics of Drosophila, edited by J. S. F. Barker, W. T. Starmer and R. J. MacIntyre. Plenum, New York.

Oakeshott, J. G., M. Saad, A. Y. Game & M. J. Healy, 1994. Causes and consequences of esterase 6 enzyme activity variation in pre-adult *Drosophila melanogaster* (submitted for publication).

Oakeshott, J. G., S. R. Wilson & W. R. Knibb, 1988. Selection affecting enzyme polymorphisms in confined populations of *Drosophila* living in a natural environment. Proc. Natl. Acad. Sci. USA 85: 293-297.

O'Brien, S. J., 1990. Genetic Maps: Locus maps of Complex Genomes. Cold Spring Harbor Laboratory Press, New York.

Olson, P. F., L. I. Fessler, R. E. Nelson, R. E. Sterne, A. G. Campbell & J. H. Fessler, 1990. Glutactin, a novel *Drosophila* basement membrane-related glycoprotein with a sequence similarity to serine esterases. EMBO J. 9: 1219-1227.

Okada, Y. & K. Wakabayashi, 1988. Purification and characterisation of esterases D-1 and D-2 from human erythrocytes. Arch. Biochem. Biophys. 263: 130-136.

Ollis, D. L., E. Cheah, M. Cygler, B. Dijkstra, F. Frolow, S. M. Franken, M. Harel, S. J. Remington, I. Silman, J. Schrag, J. L. Sussman, K. H. G. Verschueren & A. Goldman, 1992. The α/β hydrolase fold. Protein Eng. 5: 197-211.

Oppenoorth, F. J. & K. van Asperen, 1960. Allelic genes in the housefly producing modified enzymes that cause organophosphate resistance. Science 132: 298-299.

Ounissi, H. & P. Courvalin, 1985. Nucleotide sequence of the gene ereA encoding the erythromycin esterase in *Eschericia coli*. Gene 35: 271-278.

Parker, A. G., R. J. Russell, A. C. Delves & J. G. Oakeshott, 1991. Biochemistry and physiology of esterases in organophosphate-susceptible and -resistant strains of the Australian Sheep Blowfly, *Lucilia cuprina*. Pestic. Biochem. Physiol. 41: 305-318.

Pen, J., H. A. H. Rongen & J. J. Beintema, 1984. Purification and properties of esterase 4 from *Drosophila mojavensis*. Biochim. Biophys. Acta. 789: 203-209.

Pen, J., J. van Beeumen & J. J. Beintema, 1986. Structural comparison of two esterases from *Drosophila mojavensis* isolated by immunoaffinity chromatography. Biochem. J. 278: 691-699.

Piovant, M., I. Darboux, Y. Barthalay & R. Hipeau-Jacquotte, 1993. *Drosophila* neurotactin, an heterophilic cell recognition molecule related to cholinesterases: structural and functional characterisation of the extracellular domain, p. 74. Proc. XXth North Amer. Dros. Conf. (abstract).

Procunier, W. S., J. J. Smith & R. C. Richmond, 1991. Physical mapping of the Esterase-6 locus of *Drosophila melanogaster*. Genetica 84: 203-208.

Raftos, D. A. & P. B. Hughes, 1986. Genetic basis of a specific resistance to malathion in the Australian sheep blowfly, *Lucilia cuprina* (*Diptera*: *Calliphoridae*). J. Econ. Entomol. 79: 553-557.

Ray, J., J. Knapp, D. Grierson, C. Bird & W. Schuch, 1988. Identification and sequence determination of a cDNA clone for tomato pectin esterase. Eur. J. Biochem. 174: 119-124.

Raymer, G., J. M. A. Willard & J. L. Schottel, 1990. Cloning, sequencing and regulation of expression of an extracellular esterase gene from the plant pathogen *Streptomyces scabies*. J. Bacteriol 172: 7020-7026.

Richmond, R. C., K. M. Nielsen, J. P. Brady & E. M. Snella, 1990. Physiology, biochemistry and molecular biology of the *Est6* locus in *Drosophila melanogaster*, pp. 273-292 in Ecological and Evolutionary Genetics of *Drosophila*, edited by J.S.F. Barker, W.T. Starmer and R.J. MacIntyre. Plenum, New York.

Rippol, D. R., C. H. Faerman, P. H. Axelsen, I. Silman & J. L. Sussman, 1993. An electrostatic mechanism for substrate guidance down the aromatic gorge of acetylcholinesterase. Proc. Natl. Acad. Sci. USA 90: 5128-5132.

Ruddle, F. H. & L. Harrington, 1967. Tissue specific isozymes of the mouse (*Mus Musculus*). J. Exp. Zool. 166: 51-64.

Russell, R. J., M. M. Dumancic, G.G. Foster, G.L. Weller, M.J. Healy & J.G. Oakeshott, 1990. Insecticide resistance as a model system for studying molecular evolution, pp. 293-314 in Ecological and Evolutionary Genetics of *Drosophila*, edited by J.S.F. Barker, W.T. Starmer and R.J. MacIntyre. Plenum, New York.

Saad, M., A. Y. Game, M. J. Healy & J. G. Oakeshott, 1994. Associations of esterase 6 allozyme and activity variation with reproductive fitness in *Drosophila melanogaster* (submitted for publication).

Saitou, N. & M. Nei, 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 4: 406-425.

Sander, D. & R. Schneider, 1991. Database of homology-derived protein structures. Proteins 9: 56-68.

Sasaki, F., 1974. Esterase isozymes of *Drosophila virilis*: developmental aspects of the α- and β-esterases and their distribution in the body parts. Japan. J. Genet. 49: 223-232.

Sasaki, F., 1975. Immunological comparison of the α- and β-esterases in *Drosophila* species. Japan. J. Genet. 50: 217-233.

Sasaki, M. & S. Narise, 1978. Molecular weight of esterase isozymes of *Drosophila* species. Dros. Inform. Serv. 53: 123-124.

Schafer, D. J., D. K. Fredline, W. R. Knibb, M. M. Green & J. S. F. Barker, 1993. Genetics and linkage mapping of *Droso-*

phila buzzatii. Heredity (in press).

Schrag, J. D. & M. Cygler, 1993. The 1.8 Å refined structure of lipase from *Geotrichum candidum*. J. Mol. Biol. 230: 575-591.

Sergeev, P. V., G. N. Yenikolopov, N. I. Peunova, B. A. Kuzin, R. A. Khechumian, L. I. Korochkin & G. P. Georgiev, 1993. Regulation of tissue specific expression of the esterase S gene in *Drosophila virilis*. Nucl. Acids Res 21: 3545-3551.

Smyth, K. A., V. K. Walker, J. G. Oakeshott & R. J. Russell, 1994. Biochemical and genetic characterisation of malathion resistance in *Lucilia cuprina* (in preparation).

Spackman, M. E., J. G. Oakeshott, K.-A. Smyth, K. M. Medveczky & R. J. Russell, 1993. A cluster of esterase genes on chromosome 3R of *Drosophila melanogaster* includes homologues of esterase genes conferring insecticide resistance in several *Diptera* (submitted for publication).

Spierer, P., A. Spierer, W. Bender & D. S. Hogness, 1983. Molecular mapping of genetic and chromomeric units in *Drosophila melanogaster*. J. Mol. Biol. 168: 35-50.

Sreerama, L. & P. S. Veerabhadrappa, 1991. Purification and properties of carboxylesterases from the termite *Odentotermes horni*. W. Insect Biochem. 21: 833-844.

Sussman, J. S., M. Harel, F. Frolor, C. Oefner, A. Goldman, L. Toker & I. Silman, 1991. Atomic structure of acetylcholinesterase from *Torpedo californica*: a prototypic acetylcholine-binding protein. Science 253: 872-879.

Taylor, W. R., 1986. The classification of amino acid conservation. J. Theor. Biol. 119: 205-218.

Triantaphyllidis, C. D. & C. Christodoulou, 1973. Studies of a homologous gene-enzyme system, esterase C, in *Drosophila melanogaster* and *Drosophila simulans*. Biochem. Genet. 8: 383-390.

Tsuno, K., N. T. Aotsuka & S. Ohba, 1984. Further genetic variation at the esterase loci of *Drosophila virilis*. Biochem. Genet. 22: 323-337.

Van der Meer, J. R., R. I. L. Eggen, A. J. B. Zehnder & W. M. de Vos, 1991. Sequence analysis of the *Pseudomonas sp.* strain P51 tcb gene cluster, which encodes metabolism of chlorinated catechols: evidence for specialisation of catechol 1,2-dioxygenases for chlorinated substrates. J. Bacteriol. 173: 2425-2434.

Van Tilbeurgh, H., M.-P. Egloff, C. Martinez, N. Rugani, R. Verger & C. Cambillau, 1993. Interfacial activation of the lipase-procolipase complex by mixed micelles revealed by X-ray crystallography. Nature 362: 814-820.

Von Deimling, O. & A. Gaa, 1992. Esterase-29 (Es29): biochemical characterisation and control by two independent gene loci of a testosterone-dependent mouse serum esterase. Biochem. Genet. 30: 421-436.

Von Deimling, O. & B. Wassmer, 1991. Genetic characterisation of esterase 28 (Es28) of the house mouse. Biochem. Genet. 29: 5-9.

Vellom, D.C., Z. Radi'c, Y. Li, N. A. Pickering, S. Camp & P. Taylor, 1993. Amino acid residues controlling acetylcholinesterase and butyrylcholinesterase specificity. Biochemistry 32: 12-17.

White, M. M., S. D. Mane & R. C. Richmond, 1988. Studies of esterase 6 in *Drosophila melanogaster*. XVIII. Biochemical differences between the slow and fast allozymes. Mol. Biol. Evol. 5: 41-62.

Ziegler, R., S. Whyard, A. E. R. Downe, G. R. Wyatt & V. K. Walker, 1987. General esterase, malathion carboxylesterase, and malathion resistance in *Culex tarsalis*. Pestic. Biochem. Physiol. 28: 279-285.

Zouros, E. & W. van Delden, 1982. Substrate preference polymorphism at an esterase locus of *Drosophila mojavensis*. Genetics 100: 307-314.

Zouros, E., W. van Delden, R. Odense & H. van Dijk, 1982. An esterase duplication in *Drosophila*; differences in expression of duplicate loci within and among related species. Biochem. Genet. 20: 929-942.

Zschunke, F., A. Salmassi, H. Kreipe, F. Buck, M.R. Parwaresch & H. J. Radzun, 1991. cDNA cloning and characterisation of human monocyte/macrophage serine esterase-1. Blood 78: 506-512.

Genetica 88: 11-28, 1993.

Conservation and change in structural and 5′ flanking sequences of esterase 6 in sibling *Drosophila* species

J. Karotam[1,2]*, A. C. Delves[2] & J. G. Oakeshott[1]
[1] *Molecular Biology and Physiology Section, CSIRO Division of Entomology, P.O. Box 1700, Canberra A.C.T. 2601, Australia*
[2] *Division of Biochemistry and Molecular Biology, Australian National University, P.O. Box 4, Canberra A.C.T. 2601, Australia*
* *Address for correspondence: Biology Department, Queen's University, Kingston, Ontario, K7L3N6, Canada*

Received 14 October 1992 Accepted in revised form 30 November 1992

Key words: Drosophila, esterase 6, interspecific sequence comparison, protein evolution, regulatory evolution

Abstract

Esterase 6 (*Est-6*/EST6) is the major β-carboxylesterase in *D. melanogaster* and its siblings *D. simulans* and *D. mauritiana*. It is expressed in several tissues but its major site of expression is the sperm ejaculatory duct of the adult male. Although EST6 activity affects reproductive fitness, there are high levels of electrophoretic and activity polymorphism, at least within *D. melanogaster* and *D. simulans*. Here we present the nucleotide sequences of an *Est-6* allele and its flanking regions from each of *D. simulans* and *D. mauritiana* and compare them with the published *D. melanogaster* sequences. As might be expected, replacement sites are significantly less divergent than exon silent sites in all comparisons, suggesting that selection is acting to maintain EST6 structure and function among the three species. Nevertheless, the ratio of the levels of replacement to silent site divergence is still much higher for *Est-6* than for seven of ten other genes (including both isozyme-coding loci) for which comparable data have been published for these species. This is consistent with the high levels of EST6 electrophoretic polymorphism within *D. melanogaster* and *D. simulans* and implies that selective constraints against amino acid change are relatively weak for EST6. By contrast, comparisons involving promotor sequences show that the level of divergence in the first 350bp 5′ of the gene is significantly lower than those for four of the six other loci for which comparable data have been published for these species. In particular, there are two perfectly conserved stretches (−1 to −158bp and −219 to −334bp) each over 100bp long included in this 350bp region. Thus the data suggest a relatively low level of selective constraint on the amino acid sequence of EST6 but a relatively high level of constraint on sequences affecting aspects of its expression.

Introduction

Interspecific nucleotide sequence comparisons are now widely used to identify regions of conservation or change in the coding sequences of genes (Lewontin, 1989). Only recently, however, have interspecific comparisons been used to assess conservation or change in flanking sequences, and in particular, promoter sequences. Comparisons of divergence rates for such sequences with those for synonymous and coding changes in the corresponding structural sequences hold great potential for addressing issues concerning the selective constraints on such regions, the neutral rate of evolutionary change and the relative importance of regulatory change in evolution. Some interspecific comparisons of 5′ flanking regions have already identified conserved elements in promoter regions (e.g. Kassis *et al.*, 1989; Jones *et al.*, 1991) and, in a few cases, have lead to the identification of specific

differences in promoter sequences responsible for interspecific variation in gene expression (Bray & Hirsh, 1986; Brennan *et al.*, 1988).

The esterase 6 gene-enzyme system of *D. melanogaster* and its sibling species *D. simulans* and *D. mauritiana* is becoming an informative model system for studying both structural and regulatory evolution. Considering first its structure, EST6 is a monomeric glycoprotein of about 60 kiloDaltons in all three species (Morton & Singh, 1985). The enzyme is electrophoretically polymorphic in the two cosmopolitan species *D. melanogaster* and *D. simulans*, and parallel latitudinal clines are found in the frequencies of the common electrophoretically slow allozyme in both species (Anderson & Oakeshott, 1984). The shared clines suggest both that this variant is subject to selection and that the molecular basis for the selection is shared in the two species.

The regulation of EST6 in these species is similar in that the major pulse of expression in all three species is in the anterior ejaculatory duct of adult males (Stein *et al.*, 1984; Morton & Singh, 1985). The enzyme is transferred to females during mating (Richmond & Senior, 1981; J. K. & J. G. O., unpubl. data) and affects the females' subsequent reproductive behavior (Scott, 1986). Various lines of evidence from *D. melanogaster* suggest that the level of EST6 expressed in the ejaculatory duct affects reproductive fitness (Gilbert *et al.*, 1981) and one restriction fragment length polymorphism about 500bp 5′ of the gene is associated with population variation in male EST6 activity levels (Game & Oakeshott, 1990). *D. melanogaster*, *D. simulans* and *D. mauritiana* are the only species in the melanogaster subgroup in which male EST6 activity is significantly higher than female activity (Richmond *et al.*, 1990), although in general *D. simulans* males have higher EST6 activity than *D. melanogaster* and *D. mauritiana* males (J. K. & J. G. O., unpubl. data).

The evidence on both the structural and regulatory features of esterase 6 thus suggests that the enzyme may have a specific function in reproduction which is subject to selection within species and conserved across *D. melanogaster*, *D. simulans* and *D. mauritiana*. As part of an ongoing investigation into the molecular bases of such selection and constraint, the present study examines patterns of divergence in both the structural and flanking sequences of the *Est-6* locus among the three species. The *D. melanogaster* gene has been cloned and characterized previously (Oakeshott *et al.*, 1987; Collet *et al.*, 1990) and the current paper reports the cloning and sequencing of the *D. simulans* and *D. mauritiana* genes. The comparison of these three sequences herein reveals marked heterogeneity in the levels of divergence along the sequenced region. In particular, while the *Est-6* coding region proves to be among the most variable of all enzyme-coding genes so far compared in these species, the first 350bp of 5′ flanking sequences prove to be among the least variable.

Materials and methods

Libraries of partial *Sau*3A digests of genomic DNA from wild type *D. simulans* and *D. mauritiana* in λEMBL4A (gift of Dr. T. Kaufman) were screened with a nick-translated probe (Sambrook *et al.*, 1989) prepared from the *Est-6* cDNA clone from *D. melanogaster* (Oakeshott *et al.*, 1987). One hybridising clone for each species was chosen for further analysis. DNA prepared from these clones was digested with restriction enzymes under conditions specified by the enzyme manufacturers. Southern blot hybridizations of these digests (Reed & Mann, 1985) were used to construct restriction maps of the clones and to localize the *Est-6* genes within these maps. Several overlapping fragments spanning the *Est-6* gene and flanking regions were cloned into plasmid vectors pTZ18U and pTZ19U (BioRad) using standard procedures (Sambrook *et al.*, 1989). Single stranded plasmid DNA was prepared by the method of Vieira & Messing (1987). Both strands of the *Est-6* genes and flanking regions were sequenced by the dideoxy chain termination method (Sanger *et al.*, 1977). Sequences were analyzed using version 6.0 of the University of Wisconsin Sequence Analysis Software Package (Devereux *et al.*, 1984).

Results

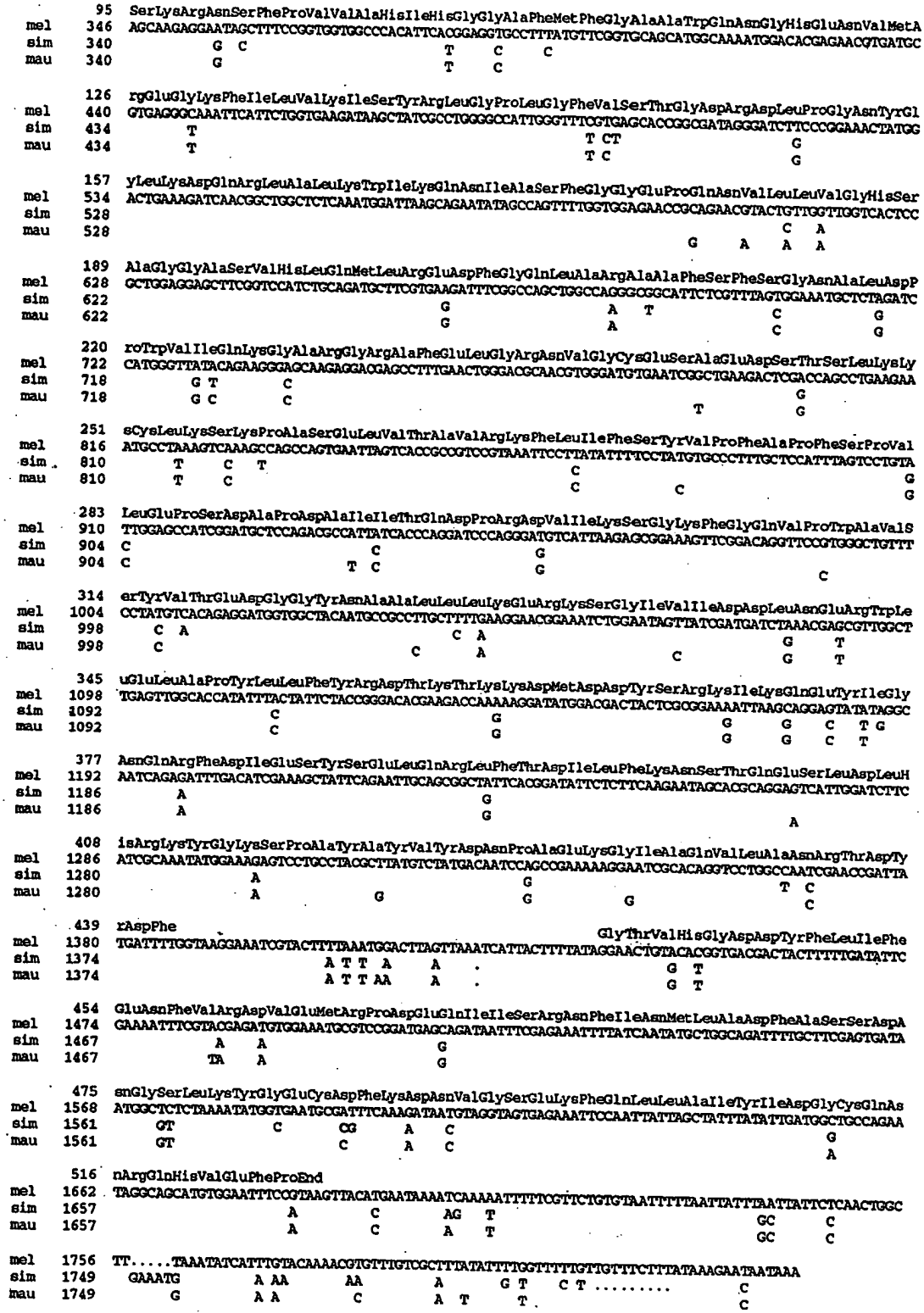
Overall structural and sequence comparisons

A 2.85kb region encompassing the *Est-6* gene has been sequenced for both *D. simulans* and *D. mauri-*

```
mel  -1131 CGAA.CTTAATGATTTTCCCTTTTTCCAAAACTTACACAAATATATACCTAATACCTTACACTGAAAAAACTGTTCAATTCACTATGTACTTAT
sim  -1137 A  TT       G    .  C T                  A       A         G A              T
mau  -1135    TT            .  C T    A             A       T           GA

mel  -1038 TAGTAGTTCTTATAAACCTTGTTATTACTAATTACGTACGAAAACAGAACGGGAAATAAATATCTAAAAATTTAAATAGTAAACTGAGCTGATG
sim  -1044     T                   AAT    C           A     G                   C              T
mau  -1039                         A T    C         G AA A                  G    C GT

mel   -944 ...TTTCTACACAAGGATGAACACAAGTATCAAAGTGCTCATAGGTACAG.....CATTTCTAGAACACAATTACACCAGTCTTTTGCTAGCAA
sim   -950 CTA       GGTCATGCTCTA AGG.....      A    T  A TAGGG      GC    G         T       A
mau   -948 ...   A   GGTCATGCTCTA TGG.....      A T        TAGGG    AGC                       A

mel   -858 TAAATTTTCACTCATCTCCCACTTTTCTCTAGAACTATCC.....TGAATTTATTTATTTGATTCTATTTTGTCGCAAGCACACTCAATTCTTG
sim   -861            G       CT                    TTTCC        A    A                    G
mau   -862   C        C   AC   CT    A               TTTCC        A    A              TG  A  G

mel   -769 ATCTTTTTTAATAATAATCGTAGCGGAAA.GCTGTATCTGTACAGGCAAGACCACCGCCCCTCAAAATGCCAACGAACTACAGTCGGATTCAAA
sim   -767         A      GT T          C G   G                        C   ..CT  G       G   C
mau   -768                GT T          C G   G                       AC   ..CT  G           C

mel   -676 AGCGTGCACCTGCAAGTGCAAGTTGTATGATATCTTGCACCATTTTATTTCGATACACGTTTTGGTACTTCGA.........TTTAATATGGC
sim   -675          C          .........      C                                A     ATTCGAATTT     ATG
mau   -676          C          .........      C                                A     ATTAGAATTT     ATG

mel   -592 TAATTGTTATTCAGAAAATAAACACAAAAACTTAGGACACATGTTCAAAAATAGAAATATGTACATATTTGTTGCAAACCTTAAATCTTATTAA
sim   -591              A      T   TTT                        GGC G   A          A                  C
mau   -592              A      T   TTT        A     .   G     GC G   A          A    G             C

mel   -498 TTA...TCTTTATCTATAATCAAAAAGCCTACTCATTTGTTAAATGGTTTCGTTTTAAAGTGCATTATCAAATTCATCTGGATAAACATTCAAA
sim   -497    ACG G                          A  A                          G              T
mau   -498    ACG G                          A  A                          G              T

mel   -407 CAAAAAATTGCATGCATGTAAAGTAATTAATCAAGTTGACTCGAAGCGATTGTGCAATATAAGTCTCACCTGAAGTGGGATTGGCAACTGCTTG
sim   -403          .          A                    T              A      A  T
mau   -404                     A                    T              A      A

mel   -313 GGATGCATGTCGAAAATCTATATTAAGCCCAGTCAAATATTTTAAGCGTAAAGTAACAAACCTAATTTGTATACGGCTATCGTTTTAATTCGCA
sim   -310                                                                                          T
mau   -310                                                                                          T

mel   -219 CACGCCTATCAACTGGATGATGTTCACACTAGAGTTACTCCCATTTCAAAGCCGGGCATTGGAAAACTAATCTCATGGCGTGCCCAGATCTCAA
sim   -216                     ...                    C          A
mau   -216                     ...                    C          A

mel   -125 TTGAGACTGGTTGACTGGTTGTTCAGGTGGCCGGGTGGCGATAAGCCGATCGATGGAATAAAAGGGGCCGCAATTGCCGCATCTCAAGATAGTT
sim   -125
mau   -125

        -21                                    MetAsnTyrValGlyLeuGlyLeuIleIleValLeuSerCysLeuTrpLeuGlySerAsnAla
mel     -31 GCGGTCTGAATTCGCCGGAGTGAGGAGCAACATGAACTACGTGGGACTGGGACTTATCATTGTGCTGAGCTGCCTTTGGCTCGGTTCGAACGCG
sim     -31                                            ......                                  C
mau     -31                                            ......                                  C

          1 SerAspThrAspAspProLeuLeuValGlnLeuProGlnGlyLysLeuArgGlyArgAspAsnGlySerTyrTyrSerTyrGluSerIleProT
mel      64 AGTGATACAGATGACCCTCTGTTGGTGCAGCTGCCCCAGGGCAAGCTACGTGGTCGCGATAATGGAAGCTACTACAGCTACGAATCGATTCCCT
sim      58       C G                                                                          G
mau      58       C G                                                                          G

         32 yrAlaGluProProThrGlyAspLeuArgPheGluAlaProGluProTyrLysGlnLysTrpSerAspIlePheAspAlaThrLysThrProVa
mel     158 ACGCCGAACCGCCCACTGGCGATCTACGATTCGAGGCTCCAGAGCCGTACAAACAAAAGTGGTCGGATATATTCGATGCCACCAAAACCCCGGT
sim     152                                                    A          C                   G
mau     152                  T                                            C     G             G

         63 lAlaCysLeuGlnTrpAspGlnPheThrProGlyAlaAsnLysLeuValGlyGluGluAspCysLeuThrValSerValTyrLysProLysAsn
mel     252 GGCGTGCCTGCAGTGGGATCAGTTCACGCCTGGGGCCAACAAATTGGTAGGAGAGGAGGATTGCCTAACCGTCAGCGTCTACAAGCCGAAGAAT
sim     246                                                          G          A
mau     246                                            G             G          A
```

Fig. 1. Comparison of the nucleotide sequences in the *Est-6* regions of the three species, with the complete *D. melanogaster* (mel) sequence given as a reference and differences in the *D. simulans* (sim) and *D. mauritiana* (mau) sequences shown below. Gaps inserted in any of the three sequences to improve the alignment are represented by dots. All numbering referred to in the text is that of the reference sequence, with the start of translation as +1. The amino acid sequence inferred for *D. melanogaster* is also included, and is numbered with the first residue of the mature peptide as +1. Note that resequencing of the 5′ region of the *D. melanogaster* clone Dm145 of Collet *et al.*, (1990) in this laboratory indicates two changes to its previously published sequence. We find a T at position −107 (corresponding to an A at 139 in the published sequence of Collet *et al.*, 1990) and a GG at −142 to −143 (corresponding to a single G at 103).

Fig. 1. Continued.



tiana (Genbank accession numbers L10670 (*D. simulans*), L10671 (*D. mauritiana*)). The region includes 1.13kb of 5′ flanking sequence, the 1.68kb coding region and 0.14kb of 3′ untranslated sequence. These sequences are aligned with the corresponding nucleotide sequence from *D. melanogaster* (Collet *et al.*, 1990 and K. M. Nielsen, C. Collet & R. C. Richmond, pers. comm., Genbank accession numbers M33780, M33781) in Figure 1 and the inferred amino acid sequences are aligned in Figure 2.

The overall structure of the *Est-6* gene and flanking regions is very similar among the three species. In *D. melanogaster* the coding region consists of two exons separated by a small intron (Collet *et al.*, 1990). The locations of initiation and termination codons and intron splice consensus sequences indicate that this structure is conserved in both *D. simulans* and *D. mauritiana*, although this has not been confirmed by cDNA or nuclease protection analysis in the latter two species. Exon I is 1381bp in length in *D. simulans* and *D. mauritiana* and 1387bp in *D. melanogaster*, exon 2 is 248bp long in all three species and the intron is 50bp in *D. simulans* and *D. mauritiana* and 51bp in *D. melanogaster*. The inferred amino acid sequences are 544 residues long in *D. melanogaster* and 542 residues long in *D. simulans* and *D. mauritiana*.

Analysis of promoter sequences is confined to the first 1.1kb 5′ of these genes because germ line transformation of *D. melanogaster* with the *Est-6* gene from each of the three species indicates that this is sufficient for most aspects of wild type *Est-6* expression (J. K., M. J. Healy, M. M. Dumancic & J. G. O., unpubl. data). Features of this region shared by the three sequences include a non-consensus TATA box (AATAAAA, −68 to −62bp), a potential GC box (GGCCGGG, −96 to −90bp), a potential CCAAT box (TCAAT, at −128 to −124bp) and seven contiguous nucleotides between −41 and −35bp shown by primer extension analysis to be the transcription initiation region in *D. melanogaster* (Collet *et al.*, 1990). The 3′ untranslated region in *D. melanogaster* contains two consensus polyadenylation sequences (AATAAA, Wickens & Stephenson, 1984), starting 9bp and 141bp 3′ of the stop codon, which produce two alternate transcripts in that species (Collet *et al.*, 1990). Consensus polyadenylation sequences are present in similar positions in the *D. simulans* and *D. mauritiana* sequences and northern analysis of adult RNA from both species suggests that they each produce two alternate transcripts of similar size to those of *D. melanogaster* (J. K. & J. G. O., unpubl. data). The second of the polyadenylation signals coincides with the presumed TATA box of the *esterase P* (*Est-P*) gene, a tandem duplication of *Est-6* in these species (Collet *et al.*, 1990; J. K. & J. G. O., unpubl. data). The coding regions of *Est-6* and *Est-P* are only 0.20kb apart in all three species.

Table 1 summarizes percentage nucleotide similarities in both coding and non-coding regions among the three species. For the purpose of these comparisons, the *Est-6* structural sequences have been divided into exon silent sites, exon replacement sites and intron sites, while the 5′ untranslated and 3′ untranslated sequences of the *Est-6* transcriptional unit are treated as separate segments.

Table 1. Percentages (± binomial standard errors) of nucleotide differences in various regions of *Est-6* (ignoring insertions/deletions) among *D. melanogaster* (mel), *D. simulans* (sim) and *D. mauritiana* (mau). n is the number of sites compared; UT, untranslated. All figures are corrected for multiple substitutions by the Jukes-Cantor method (Gojobori *et al.*, 1990). The fourth row (mel : sim + mau) compares the *D. melanogaster* sequence with the consensus of the other two.

	5′ Flank Distal sites (n=350)	5′ Flank Central sites (n=350)	5′ Flank Proximal sites (n=350)	5′ UT sites (n=41)	Exon Replacement sites (n=1269)	Exon Silent sites (n=357)	Intron sites (n=46)	3′ UT sites (n=137)
mel:sim	9.8 (1.8)	9.8 (1.8)	2.3 (0.8)	0	2.0 (0.4)	14.1 (2.1)	11.7 (5.4)	14.4 (3.5)
mel:mau	11.7 (1.9)	10.4 (1.8)	2.0 (0.7)	0	2.3 (0.4)	13.8 (2.1)	14.3 (6.0)	10.1 (2.9)
sim:mau	9.9 (1.7)	1.9 (0.7)	0.3 (0.3)	0	0.7 (0.2)	6.1 (1.4)	2.2 (2.2)	6.1 (2.2)
mel:sim + mau	5.3 (1.3)	9.1 (1.7)	2.0 (0.8)	0	1.8 (0.4)	10.2 (1.8)	11.7 (5.4)	8.5 (2.6)

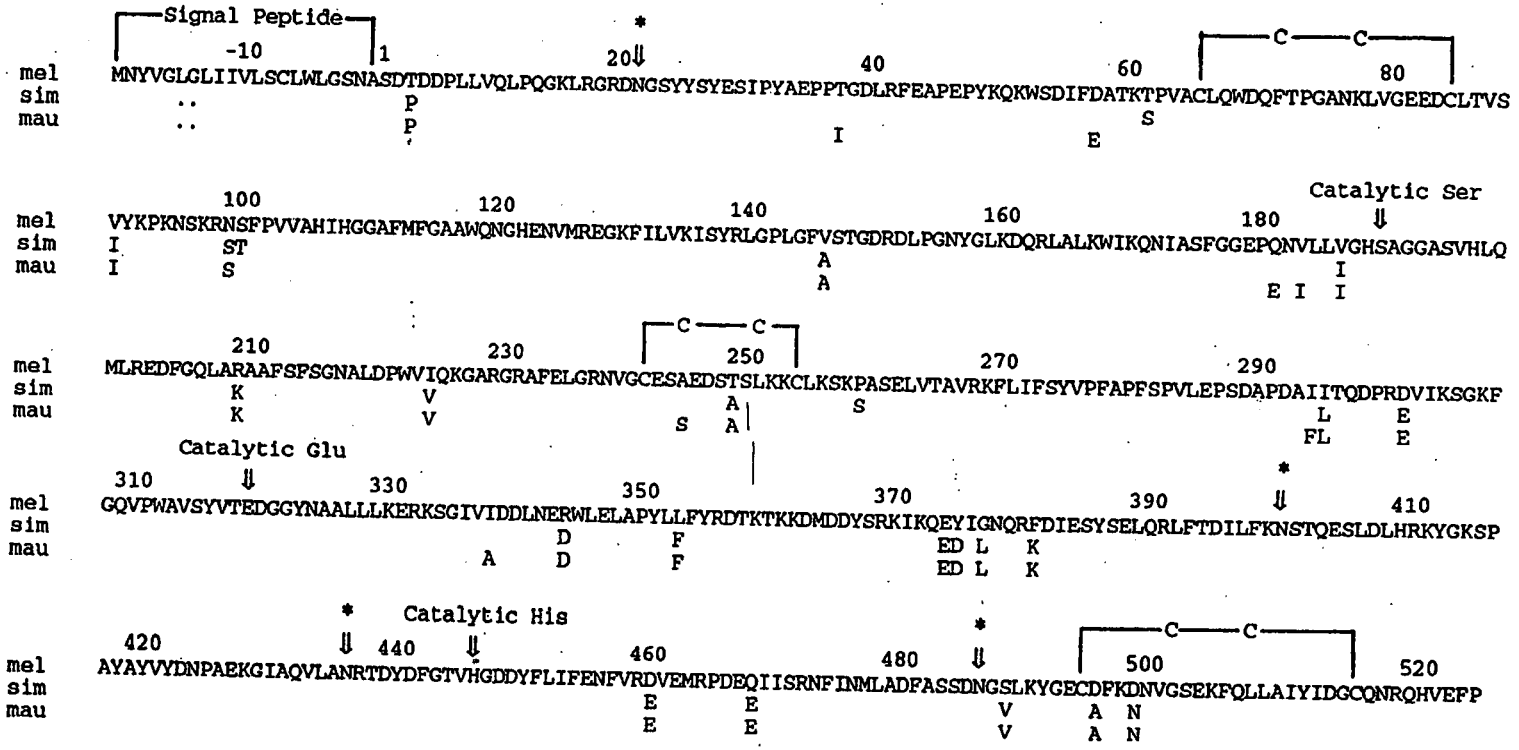


Fig. 2. Comparison of the inferred amino acid sequences of EST6 in the three species, with the complete *D. melanogaster* (mel) sequence as a reference and differences in *D. simulans* (sim) and *D. mauritiana* (mau) shown below. Gaps inserted in any of the three sequences to improve the alignment are represented by dots. Three possible catalytic residues and four N-linked glycosylation sites (*) are indicated by arrows. The signal peptide and the three disulfide bridges (C-C) are shown above the sequence. Sequences are numbered with the first residue in the mature peptide as +1.

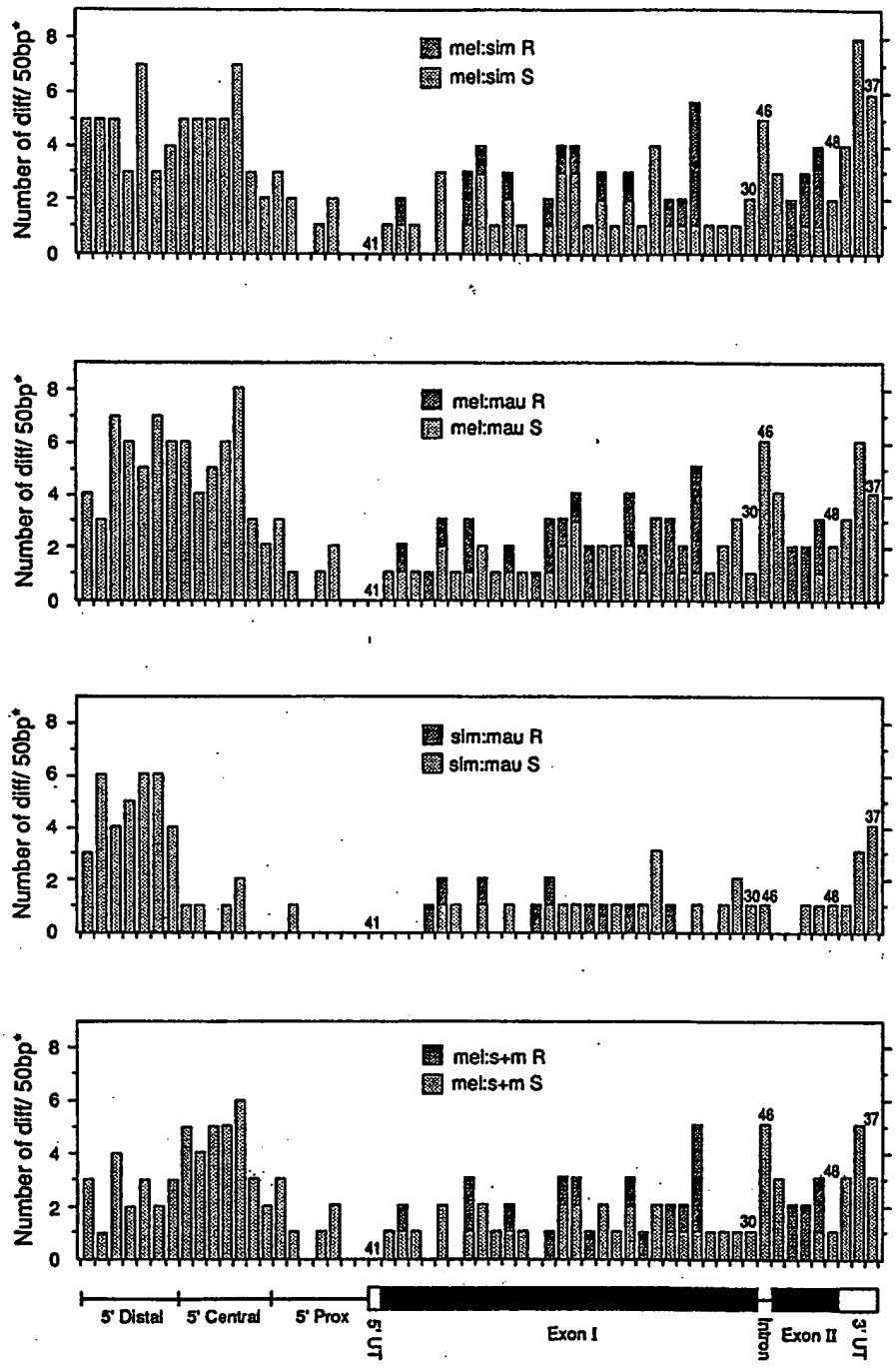


Fig. 3. Histograms showing the distribution of nucleotide differences among the sequenced regions of *Est-6* in *D. melanogaster* (mel), *D. simulans* (sim) and *D. mauritiana* (mau). The three sequences were aligned simultaneously, then divided into six regions (5′ untranscribed, 5′ untranslated, exon I, intron, exon II and 3′ untranslated sites). Each region was scored for nucleotide substitutions per 50bp interval; * where an interval contains less than 50bp, the number of sites compared is shown above the bar; R, replacement site differences; S, non-coding differences. The top three plots show the three pairwise comparisons and the fourth compares *D. melanogaster* with the consensus of the other two.

The 1050bp of untranscribed 5′ flanking sequence which can be aligned contiguously (i.e. excluding all insertion/deletions) have been divided into three equal segments: a proximal segment (−42 to −391bp), a central segment (−392 to −741bp) and a distal segment (−742 to −1091bp). Comparisons of the entire aligned sequences indicate, as is commonly accepted, that *D. simulans* and *D. mauritiana* are more closely related to each other than either is to *D. melanogaster* (Lachaise *et al.*, 1988), although the distribution of nucleotide differences along the sequenced region is higly non-random (Fig. 3).

Within the structural gene, the levels of exon silent site divergence across the three species are not significantly different to those occurring in intron sites ($G = 0.17$, $df = 1$, $P > 0.05$). Changes at silent sites and intron sites are unconstrained by selection to maintain protein structure or function, and to this extent their rates of divergence should both approximate neutral expectations (Kimura, 1991). The levels of divergence in *Est-6* replacement sites are, as expected, significantly less (5.2 fold on average) than those in silent and intron sites ($G = 77.61$, $df = 1$, $P < 0.001$), implying that replacement site change is indeed constrained by selection to maintain EST6 structure and function. Replacement site variation does not differ significantly between the two exons for any comparison ($G \leq 1.02$, $df = 1$, $P > 0.05$).

The 3′ untranslated sequences within the transcriptional unit show a similar level of divergence to the exon silent sites and the intron, but 5′ untranslated sites are invariate in all comparisons. Likewise the level of divergence among all three species in the proximal segment of the 5′ untranscribed region is significantly less (six fold on average) than that in exon silent, intron and 3′ untranslated sites ($G = 44.56$, $df = 1$, $P < 0.001$). In the central 5′ untranscribed segment, the level of divergence is lower than the silent site level for the comparison between *D. simulans* and *D. mauritiana* ($G = 7.81$, $df = 1$, $P < 0.05$) but approximates the exon silent site levels in comparisons involving *D. melanogaster* ($G = 0.19$, $df = 1$, $P > 0.05$). In the distal segment, however, the level of divergence among all three species approximates the corresponding exon silent site level ($G = 0.04$, $df = 1$, $P > 0.05$). Thus, outside the coding region, reduced divergence indicative of selective constraint is largely confined to the 5′ untranslated region and the adjacent proximal segment (−41bp to −391bp) of untranscribed DNA.

There is also a tendency for more numerous and larger insertions/deletions in the central and distal segments (six and five respectively, of up to 10bp in length, compared to a single 3bp deletion in the proximal segment; Fig. 1). Overall, however, insertions and deletions tend to be compensatory (for example, *D. simulans* and *D. mauritiana* share both a deletion of 10bp between −656 and −647 and a 10bp insertion between −604 and −603 in the reference *D. melanogaster* sequence), so the spacing of aligned sequences remains relatively constant.

Amino acid differences

Analysis of the protein coding sequences reveals a total of 31 amino acid differences among the three species, including 30 amino acid substitutions and one insertion/deletion (Fig. 2). There is no obvious clustering of the substitutions along the EST6 primary sequence, although it is noteworthy that the one insertion/deletion occurs in the signal peptide, a region which shows a high level of interspecific divergence across diverse proteins. Inferred features of the mature protein which are absolutely conserved among the three species include six cysteine residues involved in disulfide bridges (Cooke & Oakeshott, 1989) and three non-contiguous residues, Ser-188, a basic Glu-318 and an acidic His-445, that may be involved in a charge relay which donates a proton to an ester bond during ester hydrolysis (Sussman *et al.*, 1991; Schrag *et al.*, 1991 and references therein). Two alternative but now less likely candidates for the basic and acidic components of the charge relay (Asp-160 and His-408, Myers *et al.*, 1988), are also conserved among the species. There are four recognition sequences for N-linked glycosylation in the primary sequence of *D. melanogaster* EST6, all of which are known to be glycosylated in that species (Myers, 1990). Only three of these are conserved in the other two species. The fourth has been disrupted by a Ser/Val substitution at residue 487 which is common to both *D. simulans* and *D. mauritiana*. The latter two species also share a difference from *D. melanogaster* in the presumptive signal peptide. Two adjacent residues (Gly and Leu at residues −15 and −16)

in the hydrophobic core of the signal peptide in *D. melanogaster* are not found in *D. simulans* and *D. mauritiana*.

Figure 4 classifies each of the 30 amino acid substitutions as conservative or nonconservative for charge, molecular volume, polarity and hydrophobicity. Seven of these differences are conservative for all four physicochemical properties, 18 are conservative for three properties, while only four differences are nonconservative for two properties and one is nonconservative for three. In all, 83% of amino acid differences are conservative for at least three physicochemical properties. It follows that there are no major differences between the hy-

Amino acid Subst.	Res. no.	sim	mau	Charge	Molec. Volume	Polarity	Hydro-phobicity	Hydro-pathy
Thr/ Pro	3	■	■	C	C	NC	NC	SUR
Thr/ Ile	37	•	■	C	NC	NC	C	SUR
Asp/ Glu	57	•	■	C	NC	C	C	?
Val/ Ile	89	■	■	C	NC	C	C	?
Asn/ Ser	98	■	■	C	C	C	C	SUR
Ser/ Thr	99	■	•	C	C	C	NC	SUR
Val/ Ala	145	■	■	C	C	C	C	?
Gln/ Glu	180	•	■	NC	C	C	C	?
Val/ Ile	182	•	■	C	NC	C	C	?
Val/ Ile	185	■	■	C	NC	C	C	INT
Arg/ Lys	208	■	■	C	C	C	NC	?
Ile/ Val	223	■	■	C	NC	C	C	?
Ala/ Ser	243	•	■	C	C	NC	NC	SUR
Thr/ Ala	247	■	■	C	C	NC	C	SUR
Pro/ Ser	257	■	•	C	C	NC	C	?
Ile/ Phe	292	•	■	C	C	C	C	?
Ile/ Leu	293	■	■	C	C	C	C	?
Asp/ Glu	299	■	■	C	NC	C	C	SUR
Val/ Ala	336	•	■	C	C	C	C	?
Glu/ Asp	342	■	■	C	NC	C	C	SUR
Leu/ Phe	351	■	■	C	C	C	C	?
Gln/ Glu	372	■	■	NC	C	C	C	SUR
Glu/ Asp	373	■	■	C	NC	C	C	SUR
Ile/ Leu	375	■	■	C	C	C	C	SUR
Arg/ Lys	379	■	■	C	C	C	NC	SUR
Asp/ Glu	459	■	■	C	NC	C	C	?
Gln/ Glu	467	■	■	NC	C	C	C	SUR
Ser/ Val	487	■	■	C	C	NC	NC	SUR
Asp/ Ala	494	■	■	NC	C	NC	NC	SUR
Asp/ Asn	497	■	■	NC	C	C	C	SUR

Fig. 4. Comparisons of physicochemical properties of the 30 amino acid differences among the three species. The residues in the *D. melanogaster* sequence are given first, followed by the substituted residue; dots in the *D. simulans* (sim) and *D. mauritiana* (mau) columns indicate identity with *D. melanogaster*, filled boxes indicating differences. Conservative (C) and nonconservative (NC) differences are as defined by Taylor (1986). Hydropathy predictions are based on the algorithm of Kyte and Doolittle (1982): SUR, protein surface; INT, interior; ?, intermediate hydropathy (see Fig. 5).

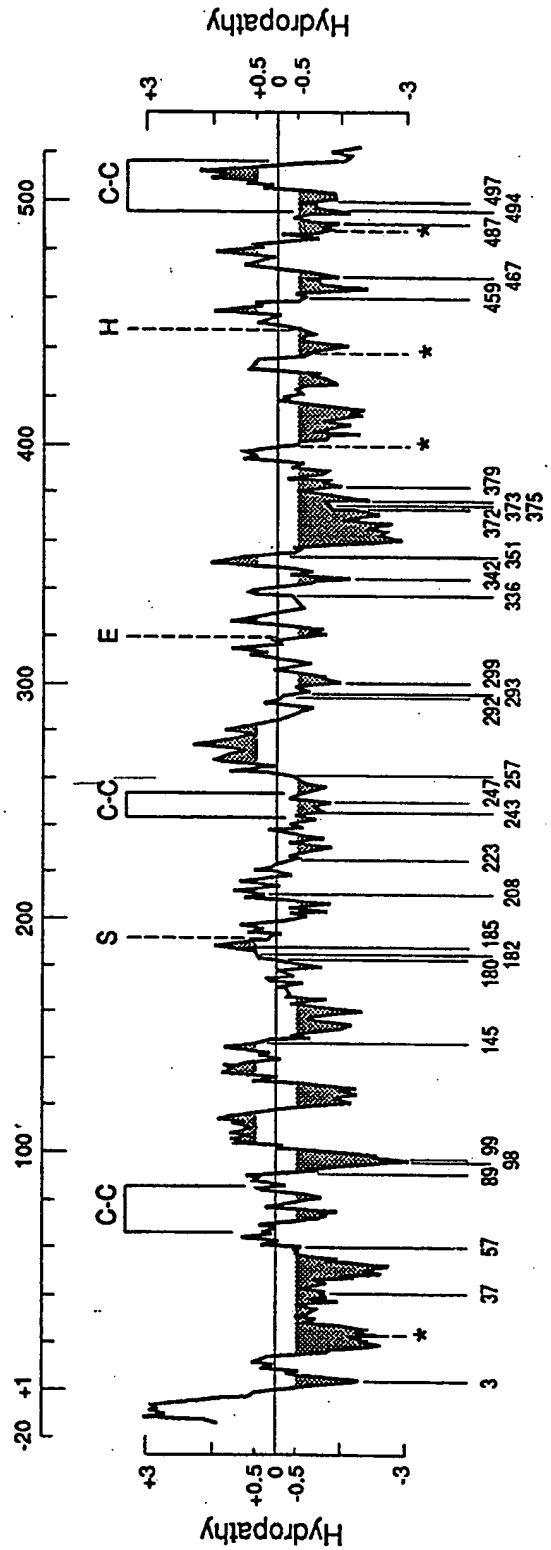


Fig. 5. Hydropathy plot for the mature *D. melanogaster* EST6 protein (negative values hydrophilic, segment size = 21), showing the positions of the 30 amino acid substitutions in *D. simulans* and *D. mauritiana*. The number line refers to the primary sequence of the protein. *, glycosylation site; S, E and H, catalytic serine, glutamic acid and histidine, respectively; C-C, disulfide bridge (see Fig. 2). Hydrophilic and hydrophobic regions are stippled, values between −0.5 and +0.5 are considered to be of intermediate hydropathy (see Table 2).

dropathy plots for EST6 in these three species (data not shown) and in Figure 5 the amino acid substitutions are superimposed on a hydropathy plot of the *D. melanogaster* sequence (Kyte & Doolittle, 1982). Of the 30 differences, sixteen are in regions predicted to be hydrophilic and likely to be on the surface of the mature protein. Thirteen of the differences are in regions of intermediate hydropathy and only one is in a region predicted to be hydrophobic. All five differences which are each nonconservative for two or more properties are in regions predicted to be hydrophilic, so on the surface of the protein where they are less likely to affect its secondary structure.

Previous studies have classified the common electrophoretically slow (EST6-S) allozymes of *D. simulans* and *D. melanogaster* as having identical mobilities (e.g. Hyyata *et al.*, 1985). However, our preliminary surveys of Australian and American populations of *D. simulans* (methods of Cooke *et al.*, 1987) indicate that the most common EST6-S allozyme has slightly less (about 2%) anodal mobility than that of the common EST6-S form of *D. melanogaster* (J. K. & J. G. O., unpubl. data). The genes sequenced herein encode a representative of the most common EST6-S allozyme of *D. simulans* and a fast (EST6-F) allozyme in *D. mauritiana* (J. K. & J. G. O., unpubl. data) while the *D. melanogaster* sequence (Collet *et al.*, 1990) encodes a representative of the most common EST6-S allozyme in natural populations (Labate *et al.*, 1989). None of the 23 amino acid differences between the *D. melanogaster* and *D. simulans* sequences result in a major net charge difference, although one or more may result in minor conformational changes responsible for the minor mobility difference between the two. The *D. mauritiana* sequence has a net negative charge with respect to the other two, due to a negatively charged Glu at residue 180 of the *D. mauritiana* sequence where both *D. melanogaster* and *D. simulans* have an uncharged Gln residue. This more negative charge of the *D. mauritiana* sequence is consistent with its more anodal electrophoretic mobility.

Promoter variation

As explained earlier, there is a highly non-random spatial distribution of nucleotide substitutions in the sequences 5′ of the *Est-6* initiation codon. Low

```
Proximal Segment - DIRECT REPEATS

mel   ACTGGTTGACTGGTTG        (-120 to -105bp)
      ::::::::::::::::
sim   ACTGGTTGACTGGTTG        (-120 to -105bp)
      ::::::::::::::::
mau   ACTGGTTGACTGGTTG        (-120 to -105bp)

Contral Sogment - PALINDROMES

mel   AAATATGTACATATTT        (-536 to -521bp)
        :::: ::::::::::
sim   gAATAeGTACeTATTt        (-535 to -520bp)
      ::::::::::::::::
mau   gAATAeGTACeTATTt        (-536 to -521bp)

mel   ActTcgAATTtaAtaT        (-610 to -597bp)
      : :::::::: : :
sim   AATTCGAATTCGAATT        (-619 to -606bp)
      :::::::::: :::::
mau   AATTCgAATTaGAATT        (-620 to -607bp)

mel   TTTaatat..(25bp)..cacaaAAA      (-603 to -563bp)
      :::::              ::  :::
sim   TTTAAATG..(25bp)..CATTTAAA      (-602 to -562bp)
      ::::::::           ::::::::
mau   TTTAAATG..(25bp)..CATTTAAA      (-603 TO -563bp)

Distal Segment - PALINDROMES

mel   AATAAATT..(36bp)..AATTTATT      (-860 to -809bp)
      ::::::::           :::::::
sim   aATAAATT..(31bp)..AATTTATa      (-863 to -807bp)
      :::: :::           ::: ::::
mau   aATAcATT..(31bp)..AATtTATa      (-864 to -808bp)

mel   TagTaaAC..(51bp)..GTacAgcA      (-952 to -896bp)
      ::::::::           ::: :: :
sim   TACTaaAC..(49bp)..GTaaAGTA      (-968 to -904bp)
      ::::  ::           ::: ::::
mau   TACTGTAC..(45bp)..GTACAGTA      (-966 to -905bp)

Distal Segment - DIRECT REPEATS

mel   GAACACAA..(34bp)..GAACACAA      (-929 to -882bp)
        :                ::::  :::
sim   tctaaagg..(34bp)..gaacgcaa      (-932 to -885bp)
      :::::  ::          ::::  :::
mau   tctaAtgg..(34bp)..gaacAcaa      (-933 to -886bp)
```

Fig. 6. Direct repeats and palindromes in the 5′ flanking sequences *Est-6* in *D. melanogaster* (mel), *D. simulans* (sim) and *D. mauritiana* (mau). Repeated or palindrome elements are underlined, colons signify interspecific nucleotide identities and lower case letters indicate intraspecific mismatches within the repeated or palindromic motifs. Numbers in parentheses indicate the position of the motif in the respective sequences. Only motifs whose elements are at least 8bp in length (allowing no mismatches) and less than 50bp apart are presented.

levels of divergence suggestive of shared functional constraints are confined to the 5′ untranslated region of the gene, the adjacent proximal segment of untranscribed DNA and, at least in *D. simulans* and *D. mauritiana*, to the next, central segment of 5′ sequence. In fact, the perfect conservation observed over the 41bp of 5′ untranslated sequence continues into the proximal 5′ untranscribed segment, which contains two perfectly conserved regions, each over 100bp long (−42bp to −158bp and −219bp to −334bp in the reference *D. melanogaster* sequence). One of the latter contains a perfect 8bp direct repeat (−120 to −113bp and −112 to −105bp) (Fig. 6) that may have a specific promoter function.

The central segment of 5′ sequence contains three different palindromic sequences, all in the region between −520 and −620bp, but none of these are perfectly conserved in all three species (Fig. 6). The first is a 16bp perfect palindrome (−536 to −521bp) in the *D. melanogaster* sequence which is in a similar position in the other two species, and is still largely intact (6/7 matches in the central 14 bases). The second, a 16bp perfect palindrome (−619 to −606bp in the *D. simulans* sequence), shows a single mismatch in *D. mauritiana* but is absent in *D. melanogaster*, due to a 10bp deletion at −604bp. The third is a 16bp perfect palindrome in *D. simulans* and *D. mauritiana* (between −602 to −595bp and −569 to −562bp in the *D. simulans* sequence) that is absent in *D. melanogaster* due to several mismatches.

The distal segment of 5′ sequence in *D. melanogaster* contains an 8bp direct repeat (−929 to −922bp and −889 to −882bp) that is not conserved in *D. simulans* or *D. mauritiana* (Fig. 6). The repeat includes the motif (ACACAA at −927 to −922bp and −887 to −882bp) which is unique to the 5′ untranscribed region and is present a further three times in the *D. melanogaster* sequence (at −1098 to −1092bp, −939 to −934bp and −571 to −566bp). Only one of these five repeats (at −1098 to −1092bp) is present in the *D. simulans* and *D. mauritiana* sequences. Two different 16bp perfect palindromes, one in *D. melanogaster* (−860 to −853 and −816 to −809bp) and the other in *D. mauritiana* (−966 to −959bp and −912 to −905bp), exhibit one or more mismatches in the other two sequences.

Discussion

Alignment of the nucleotide sequences of the *Est-6* region of *D. melanogaster*, *D. simulans* and *D. mauritiana* reveals very similar overall gene structures but highly variable levels of conservation at the nucleotide sequence level. Conservation is lowest across the exon silent sites, the intron and the 3′ untranslated region of the transcriptional unit and the most distal segment (−742 to −1091bp) of the 5′ untranscribed region. It is generally assumed that the levels of divergence in non-coding regions will approximate the neutral mutation rate (Kimura, 1991), and consistent with this assumption, the level of silent site divergence for *Est-6*, although high, falls within the range of values observed for silent sites of other genes that have been sequenced in both *D. melanogaster* and *D. simulans* (Table 2).

Nevertheless the level of exon silent site divergence does differ significantly among the genes tabulated; the level for *Est-6* is similar to that for the majority of loci listed but all of these are at least twofold greater than for *Hsc*, *Adh*, *Hsp82* and *Mtn*. The relatively high level of divergence for *Est-6* is not associated with its percentage of effectively silent sites as this statistic is relatively constant across the eleven loci compared. Although the high *Est-6* value is consistent with the proposition (Shields *et al.*, 1988; Moriyama & Gojobori, 1992) that the rate of silent site change will be greatest for loci with a low codon bias, there are enough exceptions to this proposition (e.g. *Sod*, *sal* and ci^D) to argue against any causal connection for *Est-6*. Similarly, Bulmer *et al.*, (1991) found significant levels of variation in silent site substitution rates among 58 mammalian genes which could not be attributed to variation in codon bias alone.

Compared to the consistently high levels of divergence in the silent sites, intron sites, 3′ untranslated and distal 5′ sequences, the exon replacement sites, the 5′ untranslated region of the transcriptional unit and the proximal (−42 to −391bp) and central (−393 to −741bp) 5′ untranscribed regions exhibit varying levels of constraint. This implies that shared selection pressures are acting to maintain the integrity of the latter regions across the three species.

The level of divergence in *Est-6* replacement sites is only one sixth to one ninth that in exon silent sites and the distribution and nature of the 30

Table 2. Comparison of percent nucleotide sequence divergence (corrected for multiple hits by the Jukes-Cantor method, Gojobori *et al.*, 1990) in different regions of twelve genes sequenced in both *D. melanogaster* and *D. simulans*. 5′ regions are divided into flanking (untranscribed) sites, untranslated (UT) sites and intron sites. Coding regions are divided into exon silent sites and replacement sites. The percentage of effectively silent sites (%ESS, Nei and Gojobori, 1986) and the codon bias (%G + C at third positions of codons) for each coding region are also included. Loci are listed in order of decreasing R/S value (the ratio of the percentages of replacement and exon silent site divergence). [a] O'Neil and Belote (1992); [b] Orenic *et al.* (1990) and Berry *et al.* (1991); [c] Reuter *et al.* (1989); [d] present study and Collet *et al.* (1990); [e/f] Kreitman and Hudson (1991); [g] Villares and Cabrera (1987) and Martin-Campos *et al.* (1992); [h] Lange *et al.* (1990); [i] Seto *et al.* (1987) and Kwiatowski *et al.* (1989); [j] Ingolia and Craig (1982); [k] Blackman and Meselson (1986); [l] Martin *et al.* (1988).

LOCUS	5′ Flanking Sites (n = 350)	5′ UT Sites (n)	5′ Intron Sites (n)	Exon Sites Compared	Exon Replacement Sites	Exon Silent Sites	%ESS	Codon Bias	R/S
transformer (*tra*) [a]				552	4.4	15.8	22.9	55	0.276
cubitus interruptus Dominant (ci^D) [b]	-	-	-	963	2.8	12.4	21.7	31	0.228
spalt (*sal*) [c]	-	-	-	417	2.9	17.1	23.6	29	0.169
esterase 6 (*Est-6*) [d]	2.3	0 (41)	-	1626	2.0	14.1	22.0	55	0.142
alcohol dehydrogenase-dup (*Adh-dup*) [e]	-	-	-	813	1.0	14.6	22.2	55	0.065
alcohol dehydrogenase (*Adh*) [f]	6.6	1.9 (157)	3.1 (616)	765	0.3	5.4	25.1	83	0.065
achaete (*ac*) [g]	1.7	6.6 (63)	-	603	0.8	13.1	20.8	53	0.065
metallothionein (*Mtn*) [h]	1.7	3.4 (119)	-	120	1.1		20.3	77	0
superoxide dismutase (*Sod*) [i]	12.7	3.0 (68)	-	456	0	10.9	23.8	76	0
heat shock cognate (*Hsc*) [j]	-	-	-	201	0	6.7	23.4	72	0
heat shock protein 82 (*Hsp82*) [k]	7.5	4.8 (150)	8.5 (682)	1125	0	5.6	21.1	79	0
salivary glue protein (*Sgs-3*) [l]	15.8	11.1 (29)	-	-	-	-	-	-	-

interspecific amino acid substitutions observed provides further evidence that replacement sites are selectively constrained. Firstly, hydropathy plots drawn from the three inferred amino acid sequences are almost identical, reflecting the fact that most of the interspecific amino acid differences are physicochemically conservative. Secondly, these 30 differences are not clustered in the primary sequence and are generally not found in regions known to be important to the structure or function of the mature EST6 protein. One of two possible exceptions involves the fourth recognition sequence for N-linked glycosylation in the *D. melanogaster* sequence, which is disrupted in both *D. simulans* and *D. mauritiana* by the Ser/Val difference at residue 487. However, this recognition sequence is polymorphic in a natural population of *D. melanogaster* and, even when present, is not always used (Cooke & Oakeshott, 1989; Myers, 1990); these observations suggest glycosylation at this site is not critical to EST6 function. The second possible exception is the presumptive signal peptide of *D. melanogaster*, which is two amino acid residues longer than those of *D. simulans* and *D. mauritiana*; however all three sequences remain within the accepted size range and exhibit the hydrophobic core and hydrophilic ends characteristic of signal peptides (von Heijne, 1984).

While the number and nature of the replacement site differences thus indicate a level of constraint on amino acid divergence among the three species, the degree of constraint is in fact weak relative to other loci for which comparative data are available for these species. As Table 2 shows, only three of ten other sequenced loci (*tra*, ci^D and *sal*) show a level of replacement site divergence higher than that for *Est-6* and, significantly, all three are DNA binding proteins (Reuter *et al.*, 1989; Orenic *et al.*, 1990; O'Neil & Belote, 1992). Replacement site divergence in *Est-6* is the highest of the three isozyme-

coding loci tabulated, which is consistent with the high levels of EST6 amino acid polymorphism detectable by both allozyme and DNA sequence surveys within *D. melanogaster* and *D. simulans* (Albuquerque & Napp, 1981; Cooke & Oakeshott, 1989; Labate *et al.*, 1989; J. K., T. M. Boyce & J. G. O., unpubl. data).

EST6 is electrophoretically polymorphic in both the cosmopolitan species *D. melanogaster* (Cooke & Oakeshott, 1989) and *D. simulans* (Albuquerque & Napp, 1981) but to date only two electrophoretic variants have been reported in the geographically restricted *D. mauritiana* (Gonzalez *et al.*, 1982). Latitudinal clines in the frequencies of the major EST6-F and EST6-S allozymes in *D. melanogaster* and *D. simulans* are parallel (Anderson & Oakeshott, 1984), suggesting that they may be subject to common selective forces and have a common molecular basis. Only two amino acid polymorphisms separate the major EST6-F and EST6-S variants of *D. melanogaster*; these are good candidates as the target for selection underlying the clines (Cooke & Oakeshott, 1989). One is charge-conservative (Ala (EST6-F) to Thr (EST6-S) at position 247), while the other (negatively-charged Asp (EST6-F) to uncharged Asn (EST6-S) at 237) is the only net charge difference between the two sequences, and so is likely to underlie the electrophoretic mobility difference.

What then do the sequences presented herein tell us about the EST6-F/EST6-S difference in *D. simulans*? The *D. simulans* sequence presented here is of an EST6-S variant, of similar net charge but slightly slower mobility than the major EST6-S variant of *D. melanogaster* (Collet *et al.*, 1990). The fact that the sequences of these two variants share a Thr at 247 and an uncharged Asn at 237 is consistent with the possibility that the molecular basis of the parallel clines is the same in the two species, although the sequence of an EST6-F variant of *D. simulans* is needed to confirm this.

The *D. mauritiana* sequence presented here is of an EST6-F variant of similar net charge but even faster mobility than the major EST6-F variant of *D. melanogaster* (Cooke & Oakeshott, 1989). However, the *D. mauritiana* EST6-F sequence is similar to the EST6-S variants of the other two species in that it also has a Thr at 247 and an uncharged Asn at 237. The faster mobility of the *D. mauritiana* EST6-F is best explained by the presence of a negatively charged Glu at 180, which is the only net charge difference between the *D. mauritiana* EST6-F sequence and the *D. melanogaster* and *D. simulans* EST6-S sequences (which both have an uncharged Gln at 180). Thus the faster mobility state of EST6-F in *D. mauritiana* represents a different amino acid replacement to that which distinguishes EST6-F and EST6-S in *D. melanogaster*. This difference between the species has no implications for the interpretation of the latitudinal clines above, since *D. mauritiana* is not known to share the EST6-S mobility forms found in *D. simulans* and *D. melanogaster*, and its restricted geographical range leaves no scope for the development of large scale latitudinal clines anyway.

Comparison of the nucleotide sequences of *D. melanogaster Est-6* and its homolog in *D. pseudoobscura* (*Est5B*) revealed 73.6% similarity between their inferred amino acid sequences (Brady *et al.*, 1990); the majority of the 139 amino acid differences are physicochemically conservative and there is no evidence of clustering of the nonconservative differences in the primary sequence, similar to the findings of the present study. This may simply reflect a constraint to maintain the overall structure and esteratic function of the enzyme, considering the two proteins are expressed in different tissues (Brady & Richmond, 1990) and presumably have different physiological functions.

Nevertheless, the *Est-6*/*Est5B* comparison gives the highest values for both replacement and silent site divergence of any coding region sequenced in both the melanogaster and obscura groups (Brady *et al.*, 1990). Both genes exhibit similar and very low codon bias. Assuming then that silent site divergence is largely unconstrained in both lineages, the ratio of replacement site to silent site divergence (R/S) should give an indication of the relative levels of constraint on replacement site change. R/S for *D. melanogaster* versus *D. simulans* (0.14 $\pm$ 0.04) is significantly less ($Z = 2.11$, $P < 0.05$) than that for *D. melanogaster* versus *D. pseudoobscura* (0.23 $\pm$ 0.02), implying that there is more constraint on EST6 sequences within the melanogaster group than there is between the melanogaster and obscura groups. This contrast may reflect the greater similarity in the tissue and developmental profiles of the enzyme (Aronshtam & Kuzin, 1974), and presumably therefore in its physiological function, between *D. melanogaster* and *D.*

simulans as compared to *D. pseudoobscura*. Elevated levels of ejaculatory duct expression is a feature of EST6 expression which is unique to the melanogaster complex and would suggest a qualitatively different function from *D. pseudoobscura* (Richmond *et al.*, 1990).

Comparison of the sequences 5′ of the *Est-6* gene reveals several conserved regions which may be important in the regulation of EST6 expression. The proximal segment (−42 to −391bp) displays a low level of divergence that is clearly much less than the corresponding levels for exon silent sites and indeed is not significantly different from the levels for replacement sites, making it one of the most conserved elements of the 2.83kb of sequence presented here. Consistent with this finding, sequence comparisons of *Est-6* in *D. melanogaster* with *Est5B* in *D. pseudoobscura* show that the most conserved 5′ region is the 174bp proximal to the translation start site in the reference *D. melanogaster* sequence (Brady *et al.*, 1990).

The selective constraints operating on the proximal 5′ sequence which we infer from these comparisons are borne out by functional data. Germ line transformation experiments with *D. melanogaster Est-6* show that at least the first 150bp of 5′ sequence are required for basal levels of expression throughout development (M. J. Healy, M. M. Dumancic & J. G. O., unpubl. data). Precisely which elements within this region are required for expression await more detailed functional analysis. However, a direct repeat associated with one of the two perfectly conserved regions in the proximal 5′ segment may have significance (Wingender, 1988). Evidence for this is the fact that a single element of this repeat (ACTGGTTG, repeated between −120 and −105bp in the *D. melanogaster* sequence) is in a similar position in the *D. pseudoobscura Est5B* 5′ sequence (Brady *et al.*, 1990) and is part of the longest stretch of perfect conservation among these four species (TGAG*ACTGGTTG* at −120 to −109 in the *D. pseudoobscura* sequence).

The level of divergence in the central segment of 5′ sequence (−326 to −650bp) is again not significantly different from the corresponding value for exon replacement sites for *D. simulans* versus *D. mauritiana* (Table 1, Fig. 3). However, this is not true for comparisons involving *D. melanogaster*, which show a level of divergence in this region which is not significantly different from that in silent sites. These observations imply either that there is less selective constraint on this region in *D. melanogaster* than in the other two species, or that the *D. melanogaster* sequences in this segment are evolving a different function. In view of the differences among the three species in the degree of male-specific expression (Morton & Singh, 1985; J. K. & J. G. O., unpubl. data), it is noteworthy here that functional studies show that this central segment is required for male-specific expression in *D. melanogaster* (M. J. Healy, M. M. Dumancic & J. G. O., unpubl. data).

One particular sequence within the central 5′ segment, a 16bp palindrome which is conserved across all three species (Fig. 6), is a candidate for a shared regulatory sequence. This palindrome contains an *Rsa*I site, a polymorphism for which is associated with variation in male EST6 activity in a natural population of *D. melanogaster* (Game & Oakeshott, 1990). The palindrome is also present in the 5′ flanking sequence of *Est5B* in *D. pseudoobscura*, (Brady *et al.*, 1990) and in a known regulatory region of the *Drosophila per* gene (Jackson *et al.*, 1986), both of which also show a degree of specificity for male reproductive tissue (Giebultowicz *et al.*, 1988; Oakeshott *et al.*, 1990). However, the correlation with male specificity is not invariant: the palindrome is lacking from another gene, *Gld*, which is also expressed in the male reproductive tract (Cavener, 1992 and references therein).

Levels of nucleotide sequence divergence in the distal 5′ segment approximate those in exon silent sites. This could indicate either an absence of promoter elements requiring conservation or perhaps the presence of different functional promoter elements causing differences in expression among the three species. The latter explanation is consistent with the finding from the germ line transformation experiments within *D. melanogaster* that this region is required for wild type levels of expression in a variety of tissues (M. J. Healy, M. M. Dumancic & J. G. O., unpubl. data). EST6 activity levels are highly variable both within and among these species, (Game & Oakeshott, 1990; J. K. & J. G. O., unpubl. data), although further interspecific analyses are needed to determine the contributions that various tissues make to the overall activities. It is interesting in this respect that an 8bp direct repeat in the distal segment of the *D. melanogaster* se-

quence is absent in the two other sequences; this repeat could be involved in an aspect of expression unique to *D. melanogaster*, particularly since it includes a motif (ACACAA), which is unique to the 5′ untranslated region and present five times in the *D. melanogaster* sequence, but only once in both the *D. simulans* and *D. mauritiana* sequences.

Table 2 compares the levels of divergence in the 5′ untranslated regions of the *Est-6* transcriptional unit and in the proximal segment of the untranscribed regions with other genes for which appropriate sequence data are available for both *D. melanogaster* and *D. simulans*. While the 5′ untranslated region of *Est-6* is perfectly conserved, in other genes it contains multiple substitutions and insertion/deletions. Two genes (*Adh* and *Hsp82*) have introns within the 5′ untranslated sequence and these are more variable again. Likewise the proximal 5′ segment of untranscribed DNA is at least as conserved in *Est-6* as the equivalent length of 5′ sequence for other genes. Taken together with the relatively weak constraint on the exon replacement sites for *Est-6*, this suggests that stabilising selection on the esterase-6 gene-enzyme system is relatively weak with respect to its amino acid sequence and relatively strong with respect to the regulation of its expression. We propose that this reflects the selective constraints on the sex, tissue and temporal specificity of its expression; the major pulse of expression in all three species occurring in the ejaculatory ducts of adult males.

Acknowledgements

We thank Chris Collet, Mira Dumancic, Karen Nielsen and Rollin Richmond for sharing unpublished data, Thom Boyce and Marion Healy for valuable comments on the manuscript, and Peter Cooke, Anne Game, Mark Myers, Wendy Odgers, Lis van Papenrecht, and, in particular, Robyn Russell for helpful advice and discussions.

References

Albuquerque, C. M. R. & M. Napp, 1981. Genetic variability at the esterase 6 locus in natural populations of *Drosophila simulans* in relation to environmental heterogeneity. Genetics 98: 399-407.

Anderson, P. R. & J. G. Oakeshott, 1984. Parallel geographical patterns of allozyme variation in two sibling *Drosophila* species. Nature 308: 729-731.

Aronshtam, A. A. & B. A. Kuzin, 1974. Investigation of the activity of esterase 6 in ontogenesis of *Drosophila melanogaster* and *Drosophila simulans*. Zh Obshch. Biol. 35: 926-933.

Berry, A. J., J. W. Ajioka & M. Kreitman, 1991. Lack of polymorphism on the *Drosophila* fourth chromosome resulting from selection. Genetics 129: 1111-1117.

Blackman, R. K. & M. Meselson, 1986. Interspecific nucleotide sequence comparisons used to identify regulatory and structural features of the *Drosophila* hsp82 gene. J. Mol. Biol. 188: 499-515.

Brady, J. P. & R. C. Richmond, 1990. Molecular analysis of evolutionary changes in the expression of *Drosophila* esterases. Proc. Natl. Acad. Sci. USA 87: 8217-8221.

Brady, J. P., R. C. Richmond & J. G. Oakeshott, 1990. Cloning of the esterase 5B locus from *Drosophila pseudoobscura* and comparison with its homolog in *D. melanogaster*. Mol. Biol. Evol. 7: 525-546.

Bray, S. J. & J. Hirsch, 1986. The *Drosophila virilis* dopa decarboxylase gene is developmentally regulated when integrated into *Drosophila melanogaster*. EMBO J. 5: 2305-2311.

Brennan, M. D., C.-Y. Wu & A. J. Berry, 1988. Tissue specific regulatory differences for the alcohol dehydrogenase genes of Hawaiian *Drosophila* are conserved in *Drosophila melanogaster* transformants. Proc. Natl. Acad. Sci. USA 85: 6866-6869.

Bulmer, M., K. H. Wolfe & P. M. Sharp, 1991. Synonymous nucleotide substitution rates in mammalian genes: Implications for the molecular clock and the relationship of mammalian orders. Proc. Natl. Acad. Sci. USA 88: 5974-5978.

Cavener, D. G., 1992. Transgenic animal studies on the evolution of regulatory circuitries. BioEssays 14: 237-244.

Collet, C., K. M. Nielsen, R. J. Russell, M. Karl, J. G. Oakeshott & R. C. Richmond, 1990. Molecular analysis of duplicated esterase genes in *Drosophila melanogaster*. Mol. Biol. Evol. 7: 9-28.

Cooke, P. H. & J. G. Oakeshott, 1989. Amino acid polymorphisms for esterase 6 in *Drosophila melanogaster*. Proc. Natl. Acad. Sci. USA 86: 1426-1530.

Devereux, J., P. Haeberli & O. Smithies, 1984. A comprehensive set of sequence analysis programs for the VAX. Nucl. Acids. Res. 12: 387-395.

Game, A. Y. & J. G. Oakeshott, 1990. Associations between restriction site variation and esterase 6 gene expression in *Drosophila melanogaster*. Genetics 126: 1021-1031.

Giebultowicz, J. M., R. A. Bell & R. B. Imberski, 1988. Circadian rhythm of sperm movement in the male reproductive tract of the gypsy moth *Lymantria dispar*. J. Insect Physiol. 34: 527-532.

Gilbert, D. G., R. C. Richmond & K. Sheehan, 1981. Studies of esterase 6 in *Drosophila melanogaster* V: Progeny production and sperm use in females inseminated by males carrying active or null alleles. Evolution 35: 21-37.

Gojobori, T., E. N. Moriyama & M. Kimura, 1990. Statistical methods for estimating sequence divergence. Meth. Enzymol. 183: 531-550.

Gonzalez, A. M., V. M. Cabrera, J. M. Larruga & A. Gullon, 1982. Genetic distance in the sibling species *Drosophila melanogaster*, *Drosophila simulans* and *Drosophila mauritiana*. Evolution 36: 517-522.

Hyyata, P., P. Capy, J. R. David & R. S. Singh, 1985. Enzymatic and quantitative variation in European and African populations of *Drosophila simulans*. Heredity 54: 209-217.

Ingolia, T. D. & E. A. Craig, 1982. *Drosophila* gene related to the major heat shock-induced gene is transcribed at normal temperatures and not induced by heat shock. Proc. Natl. Acad. Sci. USA 79: 525-529.

Jackson, F. R., T. A. Bargiello, S. Yun & M. W. Young, 1986. Product of per locus of *Drosophila* shares homology with proteoglycans. Nature 320: 185-188.

Jones, C. W., M. W. Dalton & L. H. Townley, 1991. Interspecific comparisons of the structure and regulation of the *Drosophila* ecdysone-inducible gene E74. Genetics 127: 535-543.

Kassis, J. A., C. Desplan, D. K. Wright & P. H. O'Farrell, 1989. Evolutionary conservation of homeodomain binding sites and other sequences upstream and within the major transcription unit of the *Drosophila* segmentation gene engrailed. Mol. Cell. Biol. 9: 4304-4311.

Kimura, M., 1991. The neutral theory of molecular evolution: A review of recent evidence. Jpn. J. Genet. 66: 367-386.

Kreitman, M. & R. R. Hudson, 1991. Inferring the evolutionary history of the Adh and Adh-dup loci in *Drosophila melanogaster* from patterns of polymorphism and divergence. Genetics 127: 565-582.

Kwiatowski, J., F. Gonzales & F. J. Ayala, 1989. *Drosophila simulans* Cu,Zn superoxide dismutase gene sequence. Nucl. Acids Res. 17: 6735.

Kyte, J. & R. F. Doolittle, 1982. A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. 157: 105-132.

Labate, J., A. Bortoli, A. Y. Game, P. H. Cooke & J. G. Oakeshott, 1989. The number and distribution of esterase 6 alleles in populations of *Drosophila melanogaster*. Heredity 63: 203-208.

Lachaise, D., M.-L. Cariou, J. R. David, F. Lemeunier, L. Tsacas & M. Ashburner, 1988. Historical biogeography of the *Drosophila melanogaster* species subgroup, pp 159-225 in Evolutionary Biology (vol 22), edited by M. K. Hecht, B. Wallace and G. T. Prance, Plenum Press, New York.

Lange, B. W., C. H. Langley & W. Stephan, 1990. Molecular of evolution *Drosophila* metallothionein genes. Genetics 126: 921-932.

Lewontin, R. C., 1989. Inferring evolutionary events from DNA coding sequence differences. Mol. Biol. Evol. 6: 15-32.

Martin, C. H., C. A. Mayeda & E. M. Meyerowitz, 1988. Evolution and expression of the sgs-3 glue gene of *Drosophila*. J. Mol. Biol. 201: 273-287.

Martin-Campos, J. M., J. M. Comeron, N. Miyashita & M. Aguade, 1992. Intraspecific and interspecific variation at the y-ac-sc region of *Drosophila simulans* and *Drosophila melanogaster*. Genetics 130: 805-816.

Morton, R. A. & R. S. Singh, 1985. Biochemical properties, homology and genetic variation of *Drosophila* 'nonspecific' esterases. Biochem. Genet. 23: 959-973.

Moriyama, E. N. & T. Gojobori, 1992. Rates of synonymous substitution and base pair composition of nuclear genes in *Drosophila*. Genetics 130: 885-864.

Myers, M. A., 1990. Structure/function relationships of the esterase 6 protein of *Drosophila melanogaster*. PhD Thesis, Australian National University.

Myers, M. A., R. C. Richmond & J. G. Oakeshott, 1988. On the origins of esterases. Mol. Biol. Evol. 5: 113-119.

Nei, M. & T. Gojobori, 1986. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. Mol. Biol. Evol. 3: 418-426.

Oakeshott, J. G., C. Collet, R. W. Phillis, K. M. Nielsen, R. J. Russell, G. K. Chambers, V. Ross & R. C. Richmond, 1987. Molecular cloning and characterization of esterase 6, a serine hydrolase of *Drosophila*. Proc. Natl. Acad. Sci. USA 84: 3359-3363.

Oakeshott, J. G., M. J. Healy & A. Y. Game, 1990. Regulatory evolution of β-carboxyl esterases in *Drosophila*, pp 359-387 in Ecological and Evolutionary Genetics of *Drosophila*, edited by J. S. F. Barker, T. W. Starmer and R. J. MacIntyre, Plenum Press, New York.

O'Neil, M. T. & J. M. Belote, 1992. Interspecific comparison of the transformer gene of *Drosophila* reveals an unusually high degree of evolutionary divergence. Genetics 131: 113-128.

Orenic, T. V., D. C. Slusarski, K. L. Kroll & R. A. Holmgren, 1990. Cloning and characterisation of the segment polarity gene cubitus interruptus Dominant of *Drosophila*. Genes Devel. 4: 1053-1067.

Reed, K. C. & D. A. Mann, 1985. Rapid transfer of DNA from agarose gels to nylon membranes. Nucl. Acids Res. 13: 7207-7221.

Reuter, D., R. Schuh & H. Jackle, 1989. The homeotic gene spalt (sal) evolved during *Drosophila* speciation. Proc. Natl. Acad. Sci. USA 86: 5483-5486.

Richmond, R. C. & A. Senior, 1981. Esterase 6 of *Drosophila melanogaster*: Kinetics of transfer to females, decay in females and male recovery. J. Insect Physiol. 27: 849-853.

Richmond, R. C., K. N. Nielsen, J. P. Brady & E. M. Snella, 1990. Physiology, biochemistry and molecular biology of the esterase 6 locus in *Drosophila melanogaster*, pp 273-292 in Ecological and Evolutionary Genetics of *Drosophila* edited by J. S. F. Barker, T. W. Starmer and R. J. MacIntyre, Plenum Press, New York.

Sambrook, J., E. F. Fritsch & T. Maniatis, 1989. Molecular Cloning: A Laboratory Manual. (2nd Edn) Cold Spring Harbor Laboratory Press, New York.

Sanger, F., S. Nicklen & A. R. Coulson, 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA 74: 5463-5467.

Scott, D., 1986. Inhibition of female *Drosophila melanogaster* remating by a seminal fluid protein (esterase 6). Evolution 40: 1084-1091.

Schrag, J. D., Y. Li, S. Wu & M. Cygler, 1991. Ser-His-Glu triad forms the catalytic site of the lipase from *Geotrichum candidum*. Nature 351: 761-764.

Seto, N. O., S. Hayashi & G. M. Tener, 1987. The sequence of the Cu-Zn superoxide dismutase gene of *Drosophila*. Nucl. Acids Res. 15: 10601.

Shields, D. C., P. M. Sharp, D. G. Higgins & F. Wright, 1988. 'Silent' sites in *Drosophila* genes are not neutral: Evidence of selection among synonymous codons. Mol. Biol. Evol. 5:

704-716.

Stein, D. P., C. S. Tepper, N. D. Able & R. C. Richmond, 1984. Studies of esterase 6 in *Drosophila melanogaster* XVI: Synthesis occurs in the male reproductive tract (anterior ejaculatory duct) and is modulated by juvenile hormone. Insect Biochem. 14: 527-532.

Sussman, J. L., M. Harel, F. Frolow, C. Oefner, A. Goldman & I. Silman, 1991. Atomic structure of acetylcholinesterase from *Torpedo californica:* A prototypic acetylcholine-binding protein? Science 253: 872-879.

Taylor, W. R., 1986. The classification of amino acid conservation. J. Theor. Biol. 119: 205-218.

Vieira J. & J. Messing, 1987. Production of single-stranded plasmid DNA. Methods Enzymol. 153: 3-11.

Villares, R. & C. V. Cabrera, 1987. The achaete-scute complex of *D. melanogaster*: Conserved domains in a subset of genes required for neurogenesis and their homology to myc. Cell 50: 415-424.

von Heijne, G., 1984. How signal sequences maintain cleavage specificity. J. Mol. Biol. 173: 243-251.

Wickens, M. & P. Stephenson, 1984. Role of the conserved AAUAAA sequence: four AAUAAA point mutants prevent messenger RNA 3′ end formation. Science 226: 1045-1051.

Wingender, E., 1988. Compilation of transcription regulating proteins. Nucl Acids Res 16: 1879-1902.

Taddei : Wonderful summary of in vitro mutagenesis exp's

RJR

Annu. Rev. Pharmacol. Toxicol. 1994. 34:281–320

THE CHOLINESTERASES: FROM GENES TO PROTEINS

Palmer Taylor and Zoran Radić[1]
Department of Pharmacology, University of California, San Diego, La Jolla, California 92093

KEY WORDS: acetylcholinesterase, butyrylcholinesterase, alternative mRNA processing, serine hydrolase, site-specific mutagenesis

INTRODUCTION

Linkages between cholinesterases and the pharmacological sciences extend back to the mid-nineteenth century when the first organophosphate was synthesized (1) and physostigmine was recognized in the western world for possessing pharmacological activity (2). However, not until Sir Henry Dale (3) delineated two components of the cholinergic nervous system was the suggestion made that physostigmine inhibited an enzyme that catalyzed the breakdown of choline esters. Dale's and later Loewi & Navratil's (4) studies established a role for acetylcholine as a labile neurotransmitter. The high turnover number of acetylcholinesterase (AChE), the specificity of its inhibitors, and the selectivity of thiocholine-metal ion interactions provided the bases for sensitive in vitro and in situ assay systems (5–7). Several cholinesterase inhibitors remain of value as medicinal agents and insecticides, but others possess the potential for insidious use as chemical warfare agents (8).

Despite this long history of study, less than a decade has passed since the primary structure of a cholinesterase was determined (9), and only in 1991 was its crystal structure solved (10). Clearly, these recent events have added a new perspective to cholinesterase research wherein all facets of gene expression become amenable to study and structure-function relationships within this family of enzymes can be approached at an atomic level

[1]Visiting Fellow from Institute for Medical Research and Occupational Health, University of Zagreb, Croatia

of resolution. This review deals primarily with the new structural information that has emerged since these developments. Not only has this structural framework added a dimension to the study of catalytic mechanisms and inhibitor specificity, but it has also enabled investigators to extend the interpretations of earlier studies where conclusions were arrived at without benefit of a structural template.

The reader should refer to other reviews for complementary or background information. Classic though somewhat dated reviews detail catalytic mechanisms (11), biochemical and catalytic properties (11, 12), and genetics of the cholinesterases (13, 14). Recently, short overviews (15, 16) and an exhaustive review (17) with a perspective on structure have been written. A recent monograph details several of the ongoing research events in the field (18).

THE CHOLINESTERASE FAMILY OF PROTEINS

The initial sequence of cholinesterase showed no global amino acid homology with any other serine hydrolases despite similarity of functional parameters and a common pentapeptide sequence around the active center serine (9). Rather, sequence identity was evident between cholinesterase and the carboxyl-terminal region of thyroglobulin (9, 19). This discovery provided the first indication that the cholinesterases defined a new family of serine hydrolases and that this gene family possessed an unexpected diversity in that non-hydrolase functions could be subserved by a common structural matrix. Soon after *Torpedo* AChE was cloned, the *Drosophila* cholinesterase gene was located from genetic studies and its sequence determined (20). This was followed by a butyrylcholinesterase (BuChE) sequence determined by amino acid sequencing (21) and by molecular cloning (22, 23). Mammalian AChEs proved more intractable, but in 1990 the mouse, bovine, and human enzyme sequences were completed (24–26). Other cholinesterase sequences, rabbit BuChE (27), rat AChE (28), *Anopheles* cholinesterase (29), and chicken AChE (30) have been reported. Distinct hydrolases from *Dictyostelium* (31, 32), *Drosophila* and other insects (33–36), the fungi *Geotrichum* and *Candida* (37), and mammals show sequence identities. Included in the mammalian group are microsomal carboxyl esterases (38, 39), lysophospholipase (40), and cholesterol esterase (41). Other proteins, while apparently not similar in primary structure, show a common folding pattern termed the α/β hydrolase fold (42). Included in this group are a wheat carboxypeptidase with a serine hydrolase mechanism (43), dienelactone hydrolase (44), and haloalkane dehalogenase (45).

In addition, members of the tactin family, glutactin and neurotactin, are homologous to the cholinesterases, but like thyroglobulin lack hydrolase activity (46, 47). No mammalian homologue of the tactins is yet known,

but in *Drosophila* tactins are believed to function in establishing contacts between heterologous cells during development. In short, a functionally eclectic family of proteins has emerged whose functional capacities extend well beyond simple hydrolase function (Figure 1). Several recent reviews have tabulated sequence identities within this family (42, 48, 49).

Since the initial AChE cloning relied on amino acid sequence to obtain oligonucleotide probes, the disulfide bond profile was established not long after in AChE (50) and BuChE (51). Labeling with radioactive DFP distinguished the catalytic serine, S_{200} (52). The histidine, H_{440}, involved in the catalytic triad was established through mutagenesis (53), but the third component in the triad, a diacidic amino acid, E_{327}, was not defined until the crystal structure was solved (10). All members of the family possess histidine in the 440 reference position, while either glutamate (as in the cholinesterases) or aspartate is found at the position corresponding to E_{327}. Corresponding residues to E_{327} and H_{440} can be found in the hydrolases of this series; however, in some cases, the alignments require liberty in gapping the residues. The three disulfide loops (50, 51) are conserved in several proteins in the family (all cholinesterases and the *Dictyostelium* proteins); others contain the amino-terminal two loops while *Culex* Est B and juvenile hormone esterase contain only the most amino-terminal loop. The third loop present in the cholinesterases, in addition to containing the histidine of the catalytic triad, functions in intersubunit contacts forming a four-helix bundle involved in subunit association (10). An additional cysteine is found very near the carboxyl-terminus that is involved in intersubunit disulfide bonds.

Intersubunit disulfide bonding occurs with identical catalytic subunits to form dimers; typically, noncovalent associations of dimers form homomeric tetramers. Heteromeric oligomers also form between the catalytic subunits and either a lipid-linked subunit or a collagen-containing subunit. These species are shown in Figure 2*A*. In mouse AChE one splicing variant does not contain a carboxyl-terminal cysteine, resulting in a monomeric enzyme species. In some cholinesterases, an eighth cysteine is found as a free sulfhydryl in variable locations. Its role in situ is unknown, but it proved invaluable for obtaining crystals of heavy metal derivatives of *Torpedo* AChE (10).

RELATIONSHIP OF PROTEIN STRUCTURE TO GENE ORGANIZATION

A comparison of protein and gene structures of the cholinesterases from different species provides additional insights into structure-function relationships. Typically, the cholinesterases have been defined as AChEs (EC 3.1.1.7) and BuChEs (EC 3.1.1.8). The latter have broad specificity with

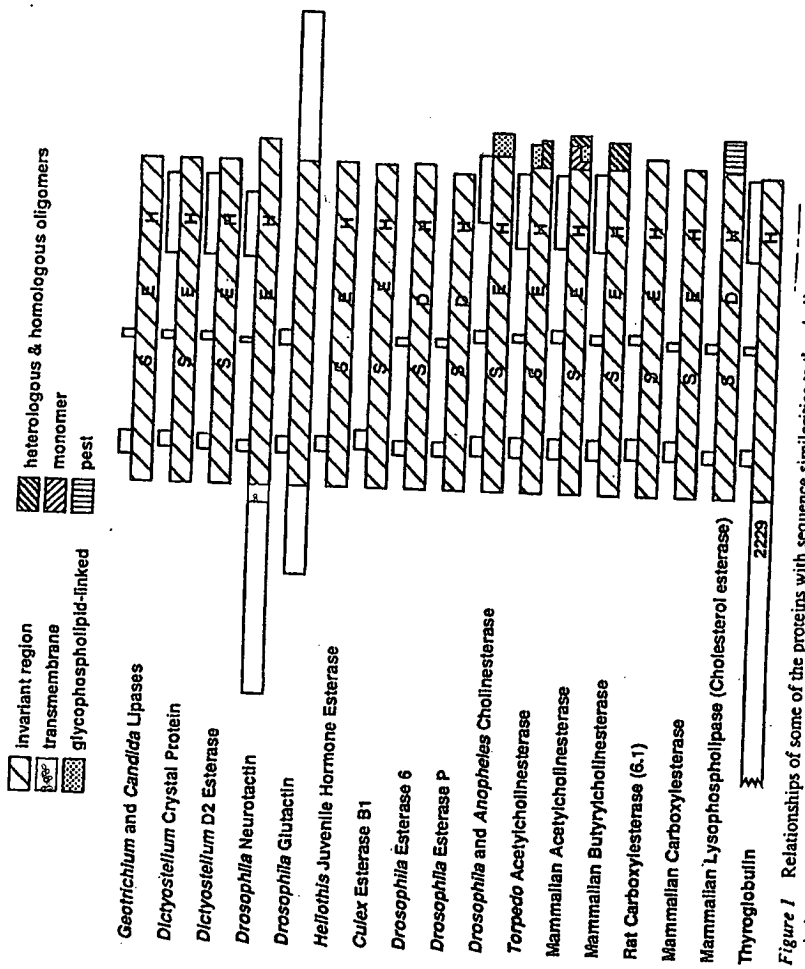


Figure 1 Relationships of some of the proteins with sequence similarities to the cholinesterases. The serines, histidines and glutamates in homologous positions to S_{200}, H_{440}, and E_{327} in *Torpedo* acetylcholinesterase are shown. Intrasubunit disulfide bonds are shown by the bracketed loops above the sequence (modified from Ref. 16).

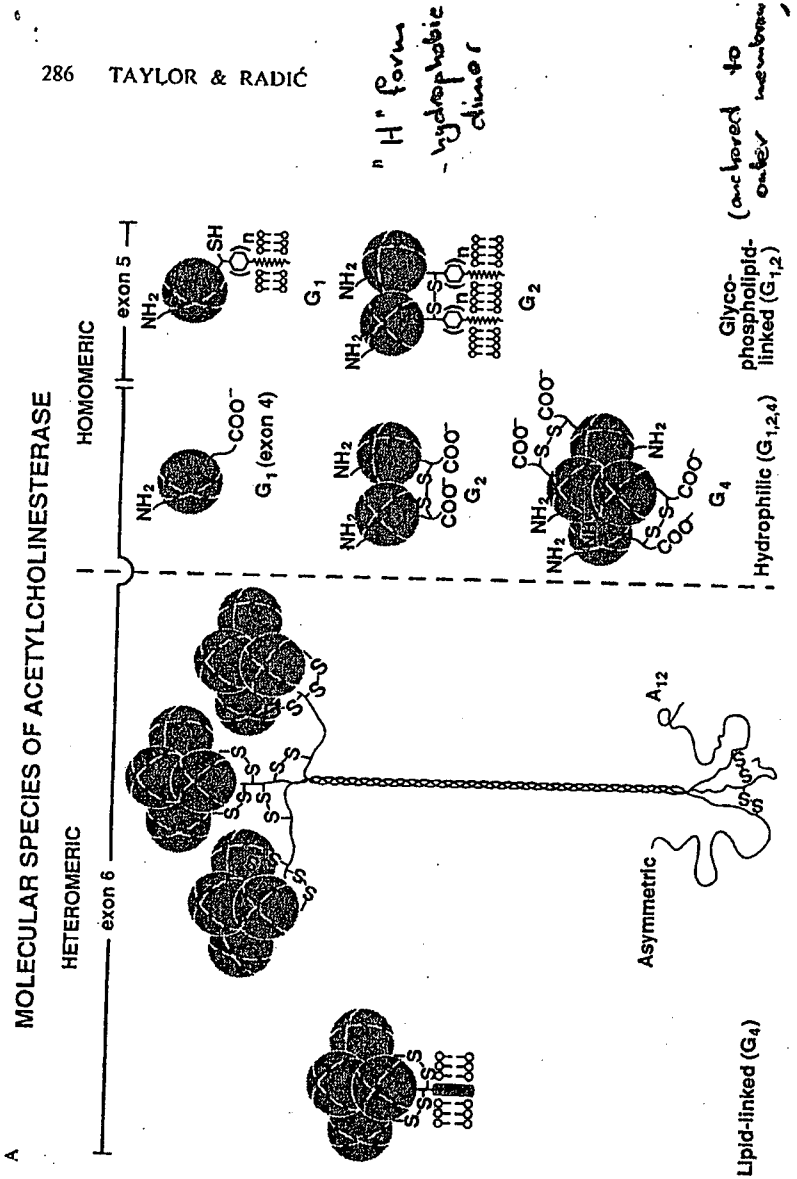
respect to the size of the substrate acyl group, while for AChE, a marked reduction in catalysis is seen between propionylcholine and butyrylcholine (54). Over the decades several selective inhibitors for AChE and BuChE have been found (55).

Drosophila appears to harbor only a single cholinesterase gene, which has features of both the AChEs and BuChEs in its encoded protein sequence (20). Similarly, its catalytic specificity is also intermediate between the two enzymes (56). Hence, it seems likely that the acetyl and butyryl subtypes of cholinesterase, which are found in lower vertebrates (57), diverged in the broad time frame between insects and lower vertebrates. Interestingly, genetic and biochemical evidence suggests multiple cholinesterase genes, perhaps three, in *Caenorhabditis elegans* (58, 59). Since the *C. elegans* genes have not been cloned, ancestral relationships in terms of sequences and specificity have yet to be ascertained.

Genomic clones of *Drosophila* cholinesterase (20), *Torpedo* AChE (60), human AChE (61), mouse AChE (61), and human BuChE (62) have been isolated. The *Drosophila* gene contains multiple exons, whereas *Torpedo* and mammalian AChE genes have relatively simple organizations. At present, our knowledge of AChE gene organization is more advanced than for BuChE, and there is as yet no evidence for alternative splicing of the BuChE gene. The open reading frame of human BuChE gene is encoded in over 50 kb of sequence and contains very large introns, whereas the comparable region in the mammalian AChE genes are encoded within 4.5–4.7 kb. The *Torpedo* AChE gene is larger; it requires 25 kb of sequence. However, the exon-intron junctions are identical in the open reading frames for AChE and BuChE, except for an additional intron located between exons 2 and 3 in mammalian AChE (Figure 2*B*).

Alternative mRNA processing is found at the 5′ and 3′ ends of the AChE gene (60, 61, 63–67), but only the splicing at the 3′ end of the open reading frame is responsible for the various molecular species of AChE. This splice occurs at amino acid 535 in the *Torpedo* sequence (68) and at 543 in mouse and human (61). Splicing in *Torpedo* gives rise to two splice alternatives, a hydrophilic peptide of 40 amino acids in length and a hydrophobic peptide of 38 amino acids; the latter appears to be cleaved after cysteine 537 with the concomitant addition of a glycophospholipid. A cDNA clone isolated from *Torpedo marmorata* has raised the possibility of a continuation of exon 4 into the retained intron (64); however, the existence of this mRNA species or the gene product awaits documentation.

In the mouse enzyme two splicing alternatives give rise to a hydrophilic species: either splicing exon 4 to exon 6 yielding a cysteine containing a 40-amino acid peptide or a direct extension into the retained intron yielding a 30-amino acid extension devoid of a cysteine (61). Hence, the latter



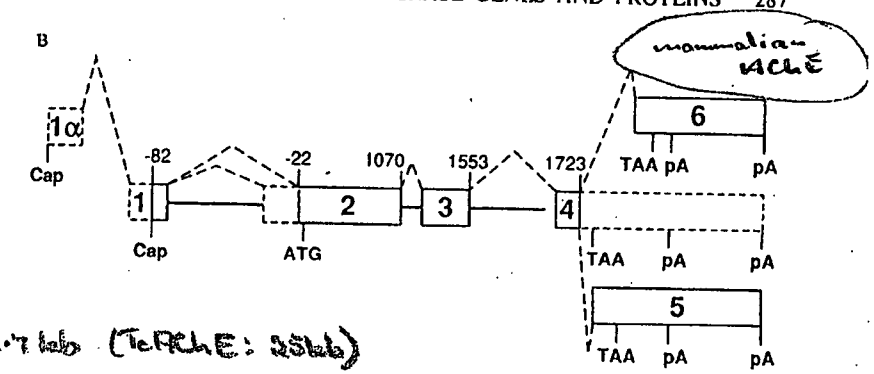


Figure 2 (*A*) Molecular species of acetylcholinesterase. The species are divided into two classes: a heteromeric class consists of catalytic subunits disulfide bonded to either a lipid-linked subunit or a triple helix of collagen-containing subunits. The homomeric class exists as monomers, dimers, and tetramers and can be divided to the hydrophilic or amphiphilic (glycophospholipid-linked) forms. The alternative exons that give rise to the various molecular species are also shown. Nomenclature designating hydrodynamic properties (A = asymmetric and G = globular) and number of catalytic subunits is also shown (17). (*B*) Structure of the genes encoding mammalian acetylcholinesterase and the exon numbering system. Alternative exon splices are shown by the dotted lines. The transcriptional start sites (Cap), translational start site (ATG), translational stop signals (TAA), and polyadenylation signals (pA) are also marked (modified from Ref. 16).

species should only exist as a monomer. The glycophospholipid-linked species in mouse and human are encoded by splicing exon 4 to exon 5 yielding 43- and 42-amino acid peptides, respectively, at their carboxyl-termini. All but 14 of the amino acids are cleaved with the addition of a glycophospholipid (24, 61). mRNA protection and expression studies verify the existence of such species in intact tissue and in transfected cells (66, 67). Hence, AChE contains a constant catalytic core consisting of the first 543 amino acids in mammals or 535 amino acids in *Torpedo*, which are encoded within three exons in mammalian AChE and two exons in *Torpedo* AChE. In mammalian BuChE the open reading frame is encoded in two exons. In this region is found the essential catalytic residues required for activity. The alternatively spliced regions in AChE only encode the remaining few amino acids (from 2 to 40) at the very carboxyl-termini of the respective processed enzymes. This domain governs intersubunit linkages and the cellular dispositions of the enzymes.

Avian AChE shows an interesting variant on this theme since it contains additional coding sequence at the position between exons 2 and 3 in the mammalian enzyme (30). The included sequence gives rise to a 20-kd increase in molecular mass of the enzyme. Variations in this region are responsible for the polymorphism of molecular weight seen in AChE from

quail (69). Alternative splicing giving rise to cholinesterases with distinct carboxyl-termini have yet to be found in the avian AChE or in BuChE from any species.

Although early studies indicated a greater complexity in the cholinesterase genes, mammalian AChE (61), avian AChE (69), and mammalian BuchE (62) are apparently each encoded by single genes. The human AChE gene is localized to 7q22 (70, 71) and human BuChE to 3q26 (71–73). The mouse gene is found at the distal end of chromosome 5, an area of synteny with 7q (74).

THREE-DIMENSIONAL STRUCTURE OF ACETYLCHOLINESTERASE

Crystallographic Analysis

The dimeric, glycophospholipid-linked form of *Torpedo* AChE was treated with phosphatidylinositol-specific phospholipase C to yield a soluble form of the enzyme amenable to crystallization (75). A structure at 2.8Å resolution has been solved and crystals suitable for higher resolution studies are available (10). Three amino acids at the amino- and carboxyl-termini, the noncleaved portion of the glycophospholipid, and a very short exposed loop, residues 485–489, showed sufficient disorder to preclude detection.

The subunits contain a 12-stranded β-sheet surrounded by 14 α-helices. They are ellipsoid in shape (45 × 60 × 65Å) and associate as dimers in a four-helix bundle. A tetramer of *Electrophorus electricus* AChE has also been crystallized (76). A low resolution structure revealed a subunit arrangement of a dimer of dimers.

Identities in Folding Patterns

The structure of *Geotrichum* lipase, an enzyme homologous in sequence, became known at about the same time as that for *Torpedo* AchE (77). These two enzymes show the same folding pattern and also contain the identical positional alignments of the Glu, His, and Ser catalytic triad discussed below. A common folding pattern is seen in the cholinesterase family (10, 49), termed the α/β hydrolase fold (49); it consists of the β_1 through β_8 sheets and the connecting α-helices. Surprisingly, a serine carboxypeptidase from wheat, a dienelactone hydrolase from *Pseudomonas*, and a haloalkane dehalogenase from *Xanthobacter* also show the same folding pattern, despite the absence of sequence identity. Even with the disparities in sequence, the structures of these proteins have converged to position the catalytic triad not only in the same three-dimensional configuration but also at corresponding positions in the turns at the ends of the β-sheets and α-helices.

Modeling of Other Cholinesterase Structures

AChE and BuChE exhibit 51–54% amino acid residue identity and modeling of BuChE on the basis of the AChE structure has been carried out, yielding a virtually identical configuration of the peptide backbone (78). Conservation of the intrasubunit disulfide bond positions and the conservation of the α/β hydrolase fold, despite considerable variations in primary structure, suggest that modeling will provide a useful framework for structural studies of other proteins in the homologous series.

The Active Center and Catalytic Triad

The crystal structure established that a E_{327} H_{440} S_{200} triad with appropriate hydrogen bonding distances and alignment was at the base of a narrow gorge 20Å in depth (10). Such triads, involving a dicarboxylic amino acid withdrawing a proton from a serine through the imidazole of histidine, are characteristic of the other families of serine hydrolases. This arrangement in the cholinesterases and *Geotrichum* lipase differs from other serine hydrolases in two respects: most enzymes in the cholinesterase family use a glutamate instead of the aspartate found in the previously characterized serine hydrolases to supply the negative charge; and the steric arrangement of residues in AChE is the mirror image of the pancreatic serine hydrolases (10). Otherwise, orientation of the side chains and hydrogen bond distances show the side chains of the triads virtually superimposable in three-dimensional space.

The gorge is lined with 14 aromatic residues. Some are deep within the gorge while most others define a large aromatic patch on the wall of the gorge. Just below the rim of the gorge lies D_{72}, at the base of the gorge lies E_{199}, and deeper into the molecule lies D_{443}. Several other anionic residues are located farther from the gorge. E_{199} is the closest anionic side chain to contact distance with trimethylammonio group acetylcholine when bound. A single negative charge at the base of the gorge seems inconsistent with a rate acceleration for binding of cationic ligands ascribable to the presence of 6–9 negative charges (79, 80). However, a global analysis of surface potentials (81) and of the orientation of the molecular dipole intrinsic to AChE with respect to the active center gorge (82) predict substantial charge accelerations for cationic substrates or inhibitors entering the gorge. Various hypotheses have also been proposed regarding the role of aromatic residues in the gorge [aromatic guidance, (10)] that facilitate diffusion of the substrate to the active center. The aromaticity may also preclude the necessity of displacement of slow-exchanging water molecules at the base

of the cleft upon ligand binding and hence it could simply play a passive role. BuChE contains six fewer aromatic residues within its gorge, yet exhibits only a threefold reduction in catalytic efficiency, as measured by k_{cat}/K_m.

Crystallographic analysis of the AChE-decamethonium and AChE-edrophonium complexes (83, 84) and the positioning of the active center serine near the carbonyl carbon of acetylcholine enable one to model the bound substrate and perform experiments on energy minimization docking. Aromatic residues clearly play an important role in stabilization of the complex. The choline moiety appears to be stabilized by W_{84} and F_{330} in AChE whose orbitals lie close to the trimethylammonio surface, as defined by its van der Waal's radii. Also, the van der Waal's surfaces of choline and E_{199} are found within 1–2Å of each other.

Several considerations allow estimation of the free energy contributions stabilizing a bound quaternary group. Studies of neutral substrate interactions with AChE (85, 86), the synthesis of cage-like compounds containing aromatic residues to stabilize quaternary ammonium ligands (87), and the crystal structure of phosphorylcholine-antibody complexes (88) all point to a role for aromatic residues being in close apposition to the quaternary moiety in the stabilization of this diverse set of complexes. However, this argument can be carried too far if longer-range electrostatic forces are ignored. In fact, both electrostatic (Coulombic) and hydrophobic forces are likely to contribute to stabilization of the complex. The approach of partitioning free energy to both the electrostatic and hydrophobic force contributions to a quaternary ligand binding site was made almost a half-century ago by Pauling and colleagues when they compared energetics of binding of phenyltrialkylammonium ions to an antibody raised to quaternary ligands (89).

As we continue around the binding site for acetylcholine (Ach), the active site serine hydroxyl should be positioned close to the carbonyl carbon on Ach. In turn, the carbonyl oxygen should be stabilized through hydrogen bonding to two amide backbone hydrogens at positions 119, 121, and/or 201 (10). A clear delineation of the acyl pocket is provided by the side chains of F_{288} and F_{290} pointing inward toward the binding site. These two residues would be expected to constrain the dimensions of the acyl pocket in AChE (Figure 3).

The Peripheral Anionic Site

J.-P. Changeux proposed an allosteric mechanism of inhibition of AChE nearly 30 years ago. He examined the inhibition of steady state kinetic parameters by various inhibitors and inhibitor combinations (90). A peripheral site, which likely gives rise to allosteric inhibition, was subsequently identified by direct titrations with the fluorescent inhibitor, propidium (91; see inset for structures). Criteria such as (*a*) the inability of agents that phosphorylate the active center serine to alter propidium binding; (*b*) the capacity of reversible inhibitors such as edrophonium and N-methylacridinium, which bind at the active center, to associate with AChE simultaneously with propidium to form ternary complexes; and (*c*) the mode of propidium inhibition of AChE acylation by substrates all point to a peripheral anionic site for the binding and allosteric actions of this inhibitor (91, 92). Moreover, measurements of fluorescence energy transfer between certain fluorescent alkyl phosphonates and propidium suggest that approximately 20Å separate the excited state dipoles between the alkylphosphate donor and the propidium acceptor of resonance energy transfer (92). Labeling studies using propidium to protect labeling by a photoactive reagent, DDF (93), and direct labeling by azidopropidium (94), have identified two sets of peptides (residues 270–278 and 251–266 in *Torpedo*) that should contribute to the binding surface of the peripheral anionic site. Finally, a terpyridine platinum coor-

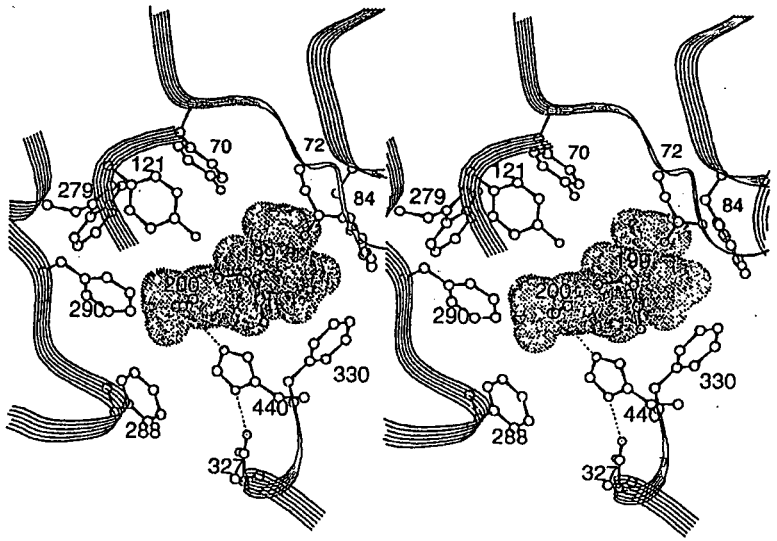
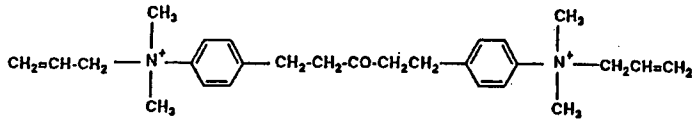
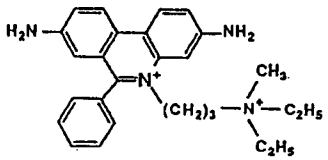


Figure 3 Structure of *Torpedo* acetylcholinesterase showing the positions of critical side chains and bound acetylcholine positioned by energy minimization (131, 132). (*a*) The catalytic triad: S_{200}, H_{440}, E_{327}. (*b*) The choline binding subsite: W_{84}, Y_{330}, E_{199}. (*c*) The acyl pocket: F_{288}, F_{290}. (*d*) The peripheral anionic site: Y_{70}, Y_{121}, W_{279}, D_{72}.

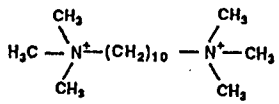
dination complex acts in a manner similar to propidium as an inhibitor and labels H_{280} in human AChE (95). The locations of the exposed surface of these residues are near the rim of the active center gorge. Hence, ligand association with the peripheral site may prevent access of substrates to the gorge by physical obstruction to restrict entry to the gorge, by charge repulsion imparted by the association of a cationic ligand, or by an allosteric mechanism in which the active center conformation is altered. In this connection, it is noteworthy that the cationic Pt-terpyridine complex inhibits catalysis of acetylcholine to a greater extent than neutral substrates (95).



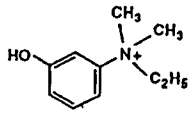
BW284C51



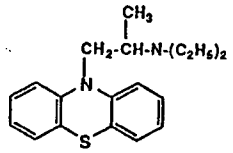
Propidium



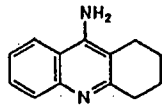
Decamethonium



Edrophonium



Ethopropazine



Tetrahydro-9-aminoacridine (Tacrine)

Three related peptide snake toxins of the fasciculin family bind to mammalian and *Torpedo* AChE but not to avian AChE or mammalian BuChE with K_Ds in the picomolar range (96, 97). These peptides of 6500 Da bind to AChE phosphorylated with DFP, but binding is prevented by propidium and certain *bis*-quaternary inhibitors. Hence, fasciculin emerges as a strong candidate for binding to the peripheral site on AChE as well (96, 97).

The function of the peripheral anionic site in catalysis in vivo and its role in synaptic activity remain open issues. It may be involved in forming

an initial complex to facilitate substrate transfer down the gorge (95). Competition between high concentrations of substrate and propidium suggest a role in substrate inhibition (98), and it has been proposed that the site serves as a sensor to maintain constant catalytic rates over a range of ionic strengths (99). Several *bis*- and *tris*-quaternary ligands bind to the peripheral site, and *bis*-quaternary ligands with large interquaternary distances (~14Å or greater) prevent the binding of both active center and peripheral site ligands (91, 92). Steric overlap between the *bis*-quaternary ligand with ligands selective for the peripheral and active sites could be responsible for this mutually exclusive binding.

Molecular Basis of Ligand Specificity at the Active Center

The dimensions of the active center gorge determined from X-ray crystallography (10) and chemical modification studies help to elucidate the specificity and orientation of bound ligands.

Early studies of Wilson & Quan (100) demonstrated the importance of a *meta* hydroxyl group in enhancing the inhibition capacity of phenyl trialkylammonium ligands. The crystal structure of the edophonium-AChE complex shows that the hydroxyl group bisects the hydrogen bond between the imidazole nitrogen in H_{440} and the serine hydroxyl group (S_{200}) and should alter the hydrogen bonding scheme (83). In addition, the aromatic ring of edrophonium is stabilized through π orbital overlap with W_{84} and, perhaps, F_{330}. The role of this site in binding of quaternary ammonium groups was also established by chemical labeling experiments where edrophonium selectively protects DDF labeling of peptides containing W_{84} in *Torpedo* (101) and presumed a peptide in *Electrophorus* AChE homologous to F_{330} (84, 102). Longer-range electrostatic interactions also appear to play a role. E_{199} resides at the base of the gorge and the distance separating the van der Waals radii of its carboxylate oxygen and the quaternary methyl groups is within 1.5Å.

Tricyclic ring-containing inhibitors such as tacrine (tetrahydro-9-aminoacridine, see inset for structures) occupy a location similar to that of edrophonium, although further rotation of the F_{330} side chain to accommodate an aromatic ring in the complex between tacrine and AChE is evident (83, 84). The tricyclic ring system inserts between F_{330} and W_{84}, causing increased stabilization by virtue of the π-orbitals. Early studies provided evidence for a charge-transfer complex between N-methylacridinium and a tryptophan in AChE (103). Moreover, the binding of N-methylacridinium and 3-aminopyridinium-1,10 decane results in near complete quenching of their fluorescence upon binding (104). The role of the indole side chain in W_{84} in acridinium binding seems clear in that it should provide the electron-rich donor ring system for association with the cation-containing

ring acceptor of acridinium. This tryptophan may well account for the changes mentioned above in absorption and fluorescence spectra typical of a charge-transfer complex.

The tricyclic ring system must not completely occlude the nucleophilic serine or the alignment of the other members of the catalytic triad since Barnett & Rosenberry found that the binding of these compounds can actually augment catalysis of neutral substrates such as ethylacetate (105). Accordingly, charge neutralization and the insertion of an aromatic ring system within the cleft enhance the catalytic surface for neutral ester substrates provided the size of the alcohol portion of the ester is kept small. Given the steric constraints of the gorge, the finding becomes even more intriguing and may argue for intrinsic flexibility within the gorge.

The portion of the active center accommodating the acyl portion of the substrate reveals that two phenylalanines, F_{288} and F_{290}, have their side chains directed into the active center and, as such, define the steric constraints of the active center. In BuChE, the conserved phenylalanines are replaced with L and I or V, providing a hydrophobic but less dimensionally constrained acyl pocket. Presumably, the phenylalanine side chains account for the marked fall-off in AChE catalysis in going from propionylcholine to butyrylcholine (54), the specificity of certain organophosphates (i.e. isoOMPA) for butyrylcholinesterase (55) and the marked stereospecificity seen with organophosphate inhibition of AChE when the moieties attached to the phosphorus differ greatly in molecular dimensions (106). Such observations would also predict that the stereoselectivity of organophosphate reactions with BuChE are much lower than with AChE.

Site of Bis-*Quaternary Ligand Association*

The site of *bis*-quaternary ligands possessing large interquaternary distances can be ascertained, in part, from kinetic studies. Early studies by Belleau and colleagues (107, 108) and by Wilson and colleagues (109) demonstrated that *bis*-quaternary and some monoquaternary inhibitors actually enhance the rate of acylation of the enzyme by neutral substrates. This enhancement is indicative of the *bis*-quaternary ligand-enzyme complex maintaining access to the active center serine for acylating agents and perhaps altering conformation of the active center to affect reactivity. In addition, series of *bis*-quaternary ligands were examined for their capacities to bind to the sulfonylated and phosphorylated AChEs (110). Only when the phosphorylating agent or the groups surrounding the ammonio group in the quaternary ion became bulky did modification of the active center serine by phosphorylation or sulfonylation affect the affinity of the *bis*-quaternary ligand (110). In addition, *bis*-quaternary ligands bind in a mutually exclusive manner with ligands selective for the active center (i.e. edrophonium and N-methylacridinium) and the peripheral site

(propidium, gallamine, and d-tubocurarine). The simplest explanation would suggest an overlap of binding surfaces. Since the interquaternary extension between the nitrogens in decamethonium is ~14Å and the two trimethylammonio groups will add another 6Å in length, the potential spanning distance is large. The crystal structure of the AChE-decamethonium complex shows one trimethylammonio group lodged between F_{330} and W_{84}; the other extends out of the active center gorge and is enlodged in the vicinity of W_{279}, Y_{70}, and Y_{121}, which reside near the lip of the gorge (83, 84). The latter residues have also been implicated in binding at the peripheral anionic site (10, 92–94). Studies with spin-labeled *bis*-quaternary ligands show immobilization of both ends of the bound molecule and a separation between the ammonio-linked nitroxides consistent with an extended bound conformation (111). Other *bis*-quaternary fluorophores have further defined the characteristics of the ligand binding site (111a).

A self-consistent picture of the binding loci of the active center, peripheral anionic site, and *bis*-quaternary ligands is emerging. Having identified the major domains in the molecule responsible for specificity, their precise roles in catalysis and in the energetics of inhibitor binding have been analyzed further through mutagenesis and molecular modeling. These studies are detailed in a subsequent section.

CATALYTIC PARAMETERS AND MECHANISMS

The catalytic potential of the cholinesterases is wide ranging with oxyesters, thioesters, selenoesters, amides, anilides, carbamoylesters, and phosphorylesters all being susceptible to catalysis (11, 12, 17, 112). Often the range of substrate catalytic potential goes unrecognized owing to the high rate of acetylcholine turnover ($k_{cat}/K_m = 10^8 M^{-1}\,sec^{-1}$) and the 10^{14} enhancement of enzyme catalyzed over H_2O catalyzed ester hydrolysis for the efficient substrates (113, 114).

A general scheme for catalysis can be represented for an ester or related substrate designated by AcOR:

$$\text{E-OH} + \text{AcOR} \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} (\text{E-OH}\cdots\text{AcOR}) \xrightarrow{k_2} \text{E-OAc} + \text{ROH} \xrightarrow[H_2O]{k_3} \text{E-OH} + \text{AcO}^- + \text{H}^+$$

acylation *deacylation*

Scheme 1

In the above scheme formation of a reversible complex with an acyl ester

is followed by acylation to form E-OAc represented by the first order rate constant k_2, and then deacylation, represented by the first order rate constant, k_3. The general features of the catalytic cycle of acylation and deacylation have been widely studied in the serine hydrolases. Serine 200 is likely to be rendered more nucleophilic by the catalytic triad. Formation of the acyl enzyme proceeds through formation of a tetrahedral intermediate which relaxes back to the trigonal, acyl enzyme. The imidazole in H_{440} may also assist by accepting the released proton. Deacylation also proceeds through a tetrahedral intermediate by attack of the acyl-enzyme bond from an internal H_2O. The H_2O may be rendered more nucleophilic by a neighboring carboxylate or imidazole residue.

In the above scheme,

$$k_{cat} = \frac{k_2\, k_3}{k_2 + k_3} \qquad \text{(Equation 1)}$$

$$K_m = \frac{k_{-1} + k_2}{k_1} \cdot \frac{k_3}{k_2 + k_3} \qquad \text{(Equation 2)}$$

$$\frac{k_{cat}}{K_m} = \frac{k_1 \cdot k_2}{k_{-1} + k_2} \qquad \text{(Equation 3)}$$

k_{cat} is governed by the energy barriers for acylation and deacylation and is the geometric mean of the two rate constants. K_m equals the equilibrium constant for the initial association only when $k_3 \gg k_2$ and $k_{-1} \gg k_2$. k_{cat}/K_m measures the initial steps leading up to formation of the acyl enzyme. Attempts to trap the acyl intermediate suggest that acylation and deacylation occur at comparable rates at V_{max} (115). This, in turn, indicates that k_2 and k_3 are of comparable magnitude for acetylcholine. For acetylcholine, k_{cat} approaches 10^4 sec^{-1} and $K_m = 5 \times 10^{-5}$ M. Accordingly, $k_{cat}/K_m = 2 \times 10^8$ M^{-1} sec^{-1}, a value approaching the diffusion limitation for k_1 (80, 114, 116).

AChE catalyzed hydrolysis of ACh and its thiol ester analogue acetylthiocholine (ATCh) approaches catalytic perfection (117) and under such conditions we might expect the transition state barriers for diffusion, acylation, and deacylation to be roughly equivalent. Hence, over a large concentration range, diffusion of substrate to the active center denoted by k_1 is essentially rate limiting.

By contrast, neutral esters and other less optimal substrates may require an induced fit to achieve acylation. Under such conditions, k_1 might be divided into two (or more) steps where k_a now reflects the diffusion step and k_b induced fit to optimize substrate orientation (118).

$$\text{E-OH} + \text{AcOR} \underset{k_{-a}}{\overset{k_a}{\rightleftharpoons}} (\text{E-OH}\cdot\cdot\text{AcOR})_1 \underset{k_{-b}}{\overset{k_b}{\rightleftharpoons}} (\text{E-OH}\cdot\cdot\text{AcOR})_2 \xrightarrow{k_2} \text{E-OAc} \xrightarrow{k_3} \underset{\text{ROH}}{\text{E-OH}} + \text{AcO}^- + \text{H}^+$$

Scheme 2

In this situation:

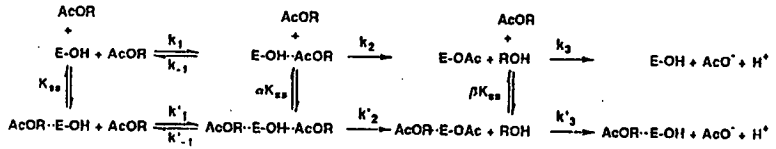
$$\frac{k_{cat}}{K_m} = \frac{k_a\, k_b\, k_2}{k_2\,(k_{-a} + k_b) + k_{-a}\, k_{-b}} \qquad \text{(Equation 4)}$$

For a common acyl group, deacylation rates should be the same; hence, we may find sets of substrates where the *a* step of diffusion of reactants or the *b* step of isomerization is rate limiting. Rosenberry (118) and Quinn and colleagues (114, 116, 119) have examined the influence of pH and fraction of deuterated substrate (isotope inventories) on catalytic parameters to deconstruct Michaelis-Menten parameters into individual rate constants and ascertain rate-limiting steps. For example, linear proton inventory plots for v/K_m have been observed for various acyl esters, which indicate that a single proton transfer rather than transfer of multiple protons is involved in the rate-limiting step of the reaction (119). Hence, no evidence can be adduced for a charge-relay system or multifunctional proton transfer in the reaction (119). The pH dependences also indicate that the rate-limiting step changes between efficient and poor substrates (118) and between ACh and benzoylcholine (116). Efficient substrates such as ACh are limited by diffusion of substrate, while others may depend either on isomerization steps leading to acylation or the acylation step itself.

In the extreme case for carbamoylating and phosphorylating agents, deacylation or the k_3 step is rate limiting in turnover. Effectively, these agents become hemisubstrates when the observation times become shorter than the deacylation half-lives.

Substrate Inhibition and Activation

Since the comprehensive studies of Augustinsson in the 1940s (54), substrate inhibition has been a hallmark of cholinesterase catalysis. It has been sufficiently characteristic to use it as a means of distinguishing AChEs from BuChEs. The mechanism of substrate inhibition is not well resolved and, in fact, data do not clearly distinguish between influence occurring on the acylation or deacylation step (98, 112, 120). If we consider the overall scheme:



Scheme 3

If excess substrate affects acylation $k_2 \neq k_2'$, while an influence on the deacylation sequence is reflected in $k_3 \neq k_3'$. For substrate inhibition either $k_2' < k_2$ or $k_3' < k_3$. Values of k_2' or $k_3' = 0$ denote excess substrate causing complete inhibition. The BuChEs (121–123) and certain mutations of AChE (123) show substrate activation. If activation and inhibition are occurring through substrate binding to a common site, we might expect both to be dependent on similar sets of residues in the molecule.

SITE-SPECIFIC MUTAGENESIS—CHOLINESTERASE CHIMERAE

Mutagenesis studies in the absence of a three-dimensional structure were largely restricted to residues where sequence conservation, sequence proximity, or a natural mutation suggested a role in function (53, 124–126). The report of a crystal structure added a new dimension as well as a flurry of activity in this arena of investigation.

Expression Systems

Initial studies of mutagenesis were done by mRNA injections into oocytes (125, 127) and transient transfections of cDNA into a receptive cell such as COS (53) or HEK cells (128). mRNA injection of single cells is labor intensive and the limited expression has not permitted a detailed analysis of kinetic and inhibition parameters. Similar limitations apply to transient transfections, particularly in the case of the *Torpedo* enzyme, where efficient protein folding does not occur at 37°C (53, 126) and expression at lower temperatures compromises cell viability. Although the high turnover rates of the cholinesterases facilitate their detection, details of substrate inhibition are only revealed at high substrate concentrations (10–100 mM). At substrate concentrations well above the K_m (~50 μM) general base catalysis of the esters will contribute substantially to basal ester hydrolysis.

Determinations of k_{cat} or k_{cat}/K_m, as measures of turnover and catalytic efficiency, necessitate titrations of stoichiometry of active sites. This entails antibody precipitation to determine total cholinesterase protein, ascertaining

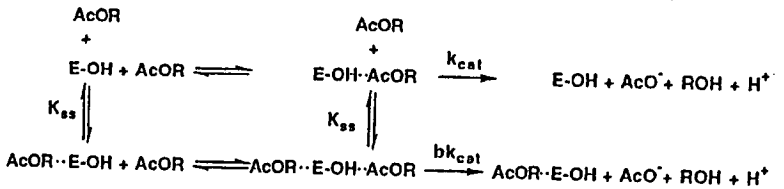
active center concentrations with high-affinity phosphorylating agents or purifying the enzyme to homogeneity. Each approach has particular advantages and limitations.

Stable transfectants of mammalian embryonic cell lines have yielded expression levels about an order of magnitude higher than transient transfection (129, 131). Finally, expression systems in baculovirus-*Spodoptera* (132) and *Escherichia coli* (133) have produced 3 mgs and over 100 mg of enzyme per liter of culture system, respectively. However, the former system presents difficult cloning steps to achieve expression, while in the latter system, generation of active enzyme required denaturation followed by refolding. Only ~3% of the enzyme renatured as an active entity.

Kinetic Parameters for Mutant Enzymes

The ratio k_{cat}/K_m, a second-order rate constant, is typically used as the measure of catalytic efficiency to compare mutant enzymes. This ratio reflects the catalytic through put under nonsaturating conditions while k_{cat} reflects maximal turnover. To describe substrate inhibition, two constants, K_m and K_{ss}, represent the concentration-dependence of catalysis and inhibition by excess substrate. An additional parameter, b, has been incorporated into kinetic schemes (112) to reflect the maximal extent of inhibition or activation by excess substrate with the mutant enzymes (123).

In a scheme where we do not differentiate whether binding of a second substrate molecule affects acylation or deacylation and $K_{ss} = \alpha K_{ss}$,



Scheme 4

then

$$v = \frac{1 + b\,[S] / K_{ss}}{1 + [S] / K_{ss}} \cdot \frac{V_{max}}{1 + K_m / [S]} \qquad \text{(Equation 5)}$$

when $b = 0$,

$$v = \frac{V_{max}}{1 + [S] / K_{ss} + K_m / [S] + K_m / K_{ss}} \qquad \text{(Equation 6)}$$

Hence, in this scheme substrate inhibition is described in terms of the dissociation constant for the inhibitory site, K_{ss}, and the relative efficiency of the ternary versus the Michaelis-Menten complex to acylate and deacylate substrate, b. This scheme is also applicable to substrate activation where $b > 1$ rather than $b <_{r} 1$. When $b = 1$, Michaelis-Menten kinetics are observed.

Reversible inhibition has been evaluated by IC_{50}'s and by measurement of dissociation constants. IC_{50}'s leave considerable uncertainties regarding inhibition mechanisms and the form of the enzyme to which the inhibitor binds. IC_{50}'s for competitive inhibitors are dependent on the K_m of the substrate relative to the substrate concentration, whereas for noncompetitive inhibitors they are independent of this ratio. Since K_m's may also be affected by mutations in the enzyme, a change in IC_{50} in the extreme case could reflect a change in K_m and not in K_I for the inhibitor. By contrast, K_I is independent of K_m. A second advantage of ascertaining the inhibition mechanism is that the influence of mutation can be compared for the same species of enzyme in the kinetic scheme. For convenience, the free species without bound substrate (i.e. E-OH) is often used; dissociation of its complex is reflected in the competitive inhibition constant.

In the case of inhibitors that carbamoylate or phosphorylate the active site serine, IC_{50}'s become parameters of limited applicability to mechanistic considerations or correlating data obtained under different conditions. Data for these inhibitors should be described in terms of a time-dependent parameter and a constant describing the concentration dependence of inhibition.

Summary of Mutation Analyses[2]

Table 1 summarizes the reported cholinesterase mutants by dividing them into several structural domains: (*a*) catalytic triad; (*b*) active center-acyl pocket; (*c*) active center-choline binding subsite; (*d*) peripheral-site(s)—rim of the gorge; (*e*) carboxyl-terminus; (*f*) glycosylation; (*g*) cholinesterase chimeras. The essential observations are detailed below:

CATALYTIC TRIAD Mutagenesis has confirmed the role for the E_{327} H_{440} S_{200} linkage in catalysis (53, 129, 134, 135). Although mutation of several other conserved diacidic amino acids results in inactive enzyme (129, 136), these residues are likely to be critical for folding into a correct tertiary conformation rather than directly involved in the acylation and deacylation steps (136). In fact, recent evidence suggests that a conformation of chicken

[2]Residue identification refers to the species under study. The parentheses refer to the *Torpedo* sequence, which serves as an alignment reference for other enzymes.

Table 1 Cholinesterase Mutations[a]

Enzyme and Residues[b]	*Torpedo* Equivalent	Catalytic, Inhibitor Specificity and Structural Change	Reference
Catalytic Triad			
TA S_{200}A,C	200	A is inactive; C may be inactive or possess 0.1% of wild-type activity	53
HA S_{203}A,C	200		129
HB S_{204}C,T,D,Q,H	200		135
TA H_{440}Q	440	Inactive; AChE with the other conserved histidine mutated is active, H_{425}	53
HA H_{447}A	440		129
TA E_{327}Q,D	327	Inactive	134
HA E_{334}D,Q,A	327	Inactive	129
Active Center-Acyl Pocket			
MA F_{295}L	288	↓ ATCh k_{cat}/K_m, ↑ BTCh k_{cat}/K_m; ↑ isoOMPA inhibition rate	131
HA F_{295}L,A	288	Similar to above	138
HB L_{286}K,Q,R,D	288	↑ K_m, change in inhibitor specificity	135
MA R_{296}S	289	Little change in activity	131
MA F_{297}I	290	↓ ATCh k_{cat}/K_m; ↓ BTCh k_{cat}/K_m; ↑ isoOMPA inhibition rate; K_{ss} ↑, b ↑	131
MA F_{295}Y	288	Little change in substrate specificity	123
MA F_{297}Y	290	↑ K_{ss}	123
DC F_{368}Y,S	290	Increased organophosphate resistance	139
HA F_{297}V,A	290	↓ ATCh k_{cat}/K_m; ↑ BTCh k_{cat}/K_m, ↑ isoOMPA inhibition	138
MA V_{300}G	293	Little change in activity	131
MA F_{295}L, F_{297}I	288, 290	↓ ATCh k_{cat}/K_m; ↑ BTCh k_{cat}/K_m; ↑ isoOMPA inhibition rate	131
HA F_{295}L, F_{297}V	288, 290	↓ ATCh k_{cat}/K_m; ↑ BTCh k_{cat}/K_m; ↑ isoOMPA inhibition	138
TA F_{288}L, F_{290}I	288, 290	↓ ATCh→↑ BTCh catalysis; ↑ isoOMPA inhibition	78
MA F_{295}L, F_{296}S, F_{297}I	288, 289, 290	↓ ATCh k_{cat}/K_m	131

Table 1 (*Continued*)

Enzyme and Residues[b]	*Torpedo* Equivalent	Catalytic, Inhibitor Specificity and Structural Change	Reference
Active Center—Choline Binding Site			
HA $Y_{337}A$	330	↓ substrate inhibition	130
MA $Y_{337}A,F$	330	Change in inhibitor specificity (esp. A)	123
TA $E_{199}Q,D$	199	↓ k_{cat}/K_m ↓ substrate inhibition (esp. D), change in inhibitor specificity, diminished aging rate	132, 53, 166
HA $E_{202}Q,D,A$	199	↓ k_{cat}/K_m ↓ substrate inhibition, change in inhibitor specificity	130
HA $W_{86}A$	84	↓ k_{cat}/K_m ATCh. ↓ propidium affinity. ↓ edrophonium affinity	130, 138
HB $Y_{440}D$	442	↑ K_m: change in inhibitor specificity	135
Gorge Entry (Peripheral Anionic Site)			
HA D_{74} E,N,G,K	72	↓ *Bis*quaternary, propidium, and dibucaine inhibition: ↓ Substrate inhibition	129, 130
MA $D_{74}N$	72	↑ K_m ↑ K_{ss}	123
HB $D_{70}G$[c]	72	Succinylcholine and dibucaine inhibition	125, 140
DC $Y_{109}D,G,K$	72	G ↑ preference BTCh K lower substrate affinity	142
TA $W_{279}A$	279	↓ Propidium and *bis*quaternary inhibition	78
HA $W_{286}A$	279	↓ Propidium and *bis*quaternary inhibition	130
MA $W_{286}R$	279	↓ Propidium and *bis*quaternary inhibition	123
MA $W_{286}A$	279	↓ Propidium and *bis*quaternary inhibition	123
MA $Y_{72}N$	70	↓ Propidium and *bis*quaternary inhibition	123
MA $Y_{124}Q$	121	↓ Propidium and *bis*quaternary inhibition	123
MA $Y_{72}N$; $Y_{124}Q$	70, 121	↓ Propidium and *bis*quaternary inhibition	123
MA $Y_{72}N$; $W_{286}R$	70, 272	↓ Propidium and *bis*quaternary inhibition	123
MA $Y_{124}Q$; $W_{286}R$	121, 279	↓ Propidium and *bis*quaternary inhibition	123
MA $Y_{72}N$; $Y_{124}Q$; W_{286}, R.A	70, 121, 279	↓ Propidium and *bis*quaternary inhibition	123
MA $Y_{72}N$; $Y_{124}Q$: $W_{286}R,A$; $D_{74}N$	see above	↓ Propidium and *bis*quaternary inhibition	123
Other Catalytic and Structural Functions			
HB $E_{441}G,E_{443}G$	443, 445	Decreased BTCh catalysis and dibucaine inhibition	140
HA $Y_{114}A$	116	Restores function to D_{70} mutants	140
HB $F_{561}Y$	563	Restores function to D_{70} mutants	140
HB $S_{425}P$[c]	427	Associated with D_{70} resistance	125, 140
HB $G_{390}V$[c]	392	↓ Succinylcholine, dibucaine and tacrine inhibition	163
HA $H_{322}N$[c]	315	YT blood group antigen	161
HA $P_{561}R$[c]	541	Allelic variation in glycophospholipid signal sequence	161
HA $F_{338}A$	331	Associates with F_{295}	138, 130
MA $F_{338}G$	331	Associates with F_{295}: ↑ K_{ss}	123
HA Y_{341}	334	↓ Substrate inhibition	138
DC $F_{115}S$[c]	78	Increased organophosphate resistance	164
DC $I_{199}V$[c]	129	Increased organophosphate resistance	164
DC $G_{303}A$[c]	227	Increased organophosphate resistance	164
Intersubunit Association			
TA C_{537}, truncation	537	Secreted	126, 165
HA $C_{580}A$	572	Secreted monomer	128
DC C_{615}, truncation	537	Secreted	145, 146, 151
Glycosylation			
HA $N_{265}Q$	258	Diminished secretion	150
HA $N_{350}Q$	343	Diminished secretion	150
HA $N_{464}Q$	457	Diminished secretion	150
HA $N_{265}Q,N_{350}Q$	258, 393	Greater diminution of secretion	150
HA $N_{256}Q,N_{464}Q$	258, 457	Greater diminution of secretion	150
HA $N_{350}Q,N_{464}Q$	343, 457	Greater diminution of secretion	150
HA $N_{765}Q,N_{350}Q,N_{464}Q$	258, 343, 457	Greater diminution of secretion	150

Table 1 (*Continued*)

Enzyme and Residues[b]	*Torpedo* Equivalent	Catalytic, Inhibitor Specificity and Structural Change	Reference
Chimerae			
TA Exon 4 deletion, exon 3–5 linkage		Glycophospholipid-linked inactive enzyme	126
HB Linkage of mutant and non-mutant enzymes	various	Augments or diminishes influence of the mutant	125, 140
MA Substituted N-terminal and/or C-terminal sequences with BuChE sequence	$B_{1-174}A_{175-575}$ $B_{1-174}A_{175-487}$ $B_{488-575}$	B_{1-174} confers BW specificity of BuChE	131
HB Substituted AChE sequence for BuChE	$B_{1-57}A_{58-133}$ $B_{134-575}$	Imparts partial AChE character	141

[a] MA = Mouse acetylcholinesterase; HA = human acetylcholinesterase; TA = *Torpedo* acetylcholinesterase; HB = human butyrylcholinesterase. DC = *Drosophila* cholinesterase

[b] Other residues, $D_{397}N$ in *Torpedo*, $D_{175}N$, $D_{404}N$ in human have been reported to produce inactive enzyme. $E_{92}Q,L$ results in inactive enzyme in *Torpedo*. Little or no change in activity was reported for $E_{84}Q$, $D_{95}N$, $D_{131}N$, $D_{333}N$, $D_{349}N$ in human and $D_{93}N$ in *Torpedo* (18)

[c] Natural mutations

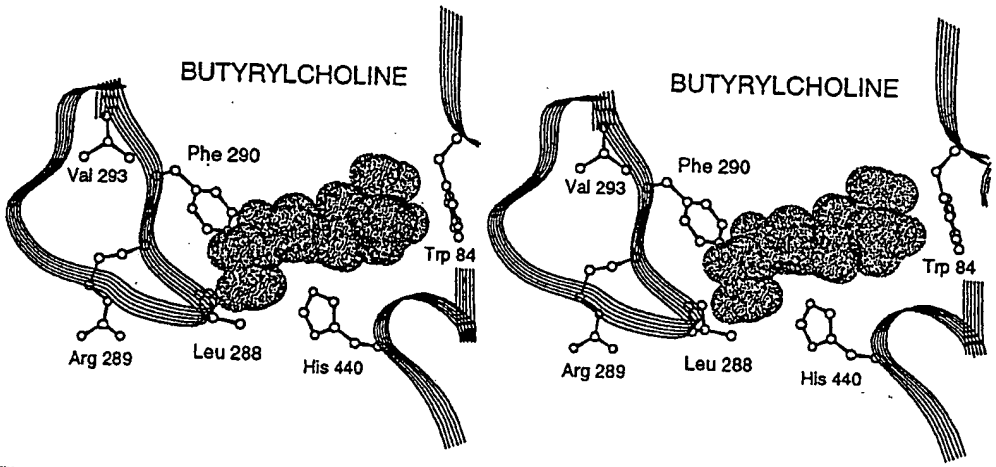


Figure 4 Structure of bound butyrylcholine within the substrate binding site of the F_{288} Mutant of Acetylcholinesterase. Energy minimization was done with the Biosym Insight II program (131).

AChE is produced that is DFP reactive, but catalytically inactive towards Ach (137). Whether the catalytically inactive mutants achieve a tertiary conformation approaching the active enzyme or are simply degraded as a nascent peptide chain is unknown. Some mutations of the active center (i.e. $S_{(200)}C$) show low activity (53, 135), and it will be of interest to achieve high expression to analyze them for catalytic properties.

The functional existence of a catalytic triad does not prove the existence of a charge-relay system or rate-limiting proton transfer (119). Rather the optimal alignments of these residues may be critical for confering a proton-withdrawing, inductive effect on the serine and/or a sink capacity for released protons.

ACTIVE CENTER ACYL POCKET Based on the residue differences between AChE and BuChE, the residues outlining the acyl pocket have been substituted in mammalian AChE to produce multiple mutant enzymes (Figure 4). Substitution of $F_{295(288)}$ and $F_{297(290)}$ in AChE to the corresponding residues found in BuChE has increased BuChE character as measured by an increased ratio of butyrylthiocholine (BTCh) to acetylthiocholine (ATCh) catalyzed hydrolysis, changes in the substrate activation and inhibition profiles, and increased susceptibility to inhibition by the BuChE-specific inhibitor, isoOMPA (131, 138). The $F_{295(288)}L$, $F_{297(290)}I$ double mutant (78, 131, 138) and the $F_{295(288)}L$, $R_{296(289)}S$, $F_{297(290)}I$ triple mutant (131) showed similar BuChE character, but were far less active. A detailed analysis of the individual F_{295} and F_{297} mutants uncovered several interesting properties of the acyl pocket. First, the $F_{295}L$ mutant of mouse AChE, while slightly less efficient towards ATCh hydrolysis, hydrolyzed BTCh with a k_{cat}/K_m greater than that found for native BuChE (131). Similar behavior was seen for the human $F_{295}L$ and $F_{295}A$ mutations (138). The $F_{297}I$ mutation is notable in its increased K_m for both ATCh and BTCh and for the elimination of substrate inhibition. In fact, the concentration dependence of BuChE catalyzed hydrolysis of BTCh and ATCh is best described in terms of substrate activation (123, 131) and the F_{297} mutation alone is sufficient to reverse the substrate inhibition in AChE and achieve a large measure of the activation seen with BuChE (123). $F_{338(331)}$, which comes in close proximity to F_{295} through ring stacking, also has a marked influence on diminishing substrate inhibition (123).

Drosophila cholinesterase has a single phenylalanine in its acyl pocket; a natural mutation to Y produces a enzyme conferring insecticide resistance to several bulky organophosphates (139).

ACTIVE CENTER-CHOLINE BINDING SUBSITE Four side chains appear to be of particular importance in stabilizing the quaternary moiety of choline. The crystal structure shows the trimethylammonio-methylene group of decamethonium or the dimethylethylammonio group of edrophonium appears to make a three-point contact with the indole ring of $W_{(84)}$ (84). $F_{(330)}$ and $Y_{(442)}$ are also in close apposition, and some movement of the side chain $F_{(330)}$ towards the aromatic ring of edrophonium is also evident in this complex. The van der Waal's outer shell of the carboxylate of $E_{(199)}$ comes within 1.5Å of that of the trimethylammonio group.

Replacement of $W_{86(84)}$ by A results in a marked reduction in ATCh catalysis and diminished binding of edrophonium (130). A follow-up study shows that the loss of activity is selective for the quaternary substrate since the isosteric, 2,2 dimethylbutyl acetate ester shows little diminution of activity (138). This finding illustrates the importance of the quaternary ammonium-indole interaction in the stabilization of complexes of substrate and inhibitors. However, a large difference in molecular volume is also inherent to this substitution. $W_{(84)}$ is conserved in all the cholinesterases.

The second aromatic residue in this domain is not conserved; the AChEs contain an F or Y at position 337(330) and BuChE has A at 332(330). Several inhibitors selective for BuChE or AChE depend on this difference. The $Y_{337}A$ mutation results in an 10- to 20-fold reduction in edrophonium affinity but little or no reduction in decamethonium affinity (123, 130, 138). By contrast, the affinities of the acridines and particularly certain phenothiazines are increased by this mutation (123). This behavior appears to depend on the phenothiazine side chain and was most marked with ethopropazine where a 2700-fold decrease in K_I was evident. This decrease was virtually identical to its difference in K_I between AChE and BuChE (123). Huperzine shows a decreased affinity with the $Y_{337}A$ substitution (A Saxena, N Qian, IM Kovach, AP Kozikowski, D Vellom, et al, submitted). Taken together, the data indicate that the aromatic group at 337(330) contributes to stabilization of the complexes (i.e. ring stacking and quaternary aromatic interactions in the case of edrophonium and stabilization of the caged structure in the case of huperzine). However, addition of the tricyclic ring system and, in particular, certain substitutions on the exocyclic chain create steric hindrance with the aromatic ring in $F_{(330)}$ or $Y_{(330)}$ This is reflected in lower affinities of inhibitors of larger volume for AChE than for either BuChE or the $Y_{337(330)}A$, AChE mutant (123). The 337(330) residue change has minimal influence on ATCh catalysis. Shafferman and colleagues have shown that the Y_{337} to A mutation diminishes substrate inhibition in human AChE (130) and suggest a direct linkage to the peripheral site. However, upon mutation of Y_{337} to A in mouse AChE substrate inhibition is still evident when examined over a wider range of substrate concentrations (123), which indicates that a substrate inhibition mechanism involving the 337 residue is not universal.

Furthermore, $Y_{(442)}$ also contributes to the choline binding site surface. In BuChE, with $F_{328(330)}$ changed to A, the role of $Y_{440(442)}$ may be more influential. Altered catalytic parameters are found with the $Y_{440}A$ mutation in BuChE (135).

Mutagenesis experiments also revealed that the charge on $E_{(199)}$ stabilizes binding in this region. Edrophonium affinity is markedly reduced (132, 130) and k_{cat}/K_m for ATCh is lowered by a factor of 50 with the $E_{199}Q$ mutation (132). Hence, the energy of stabilization of edrophonium can be partitioned between both the electrostatic and π-electron bonding forces. Similar analyses are possible for other inhibitors, substrates, and transition state mimics. The $E_{199}D$ mutation has less influence on k_{cat}/K_m but markedly affects substrate inhibition (130, 132).

THE PERIPHERAL ANIONIC SITE: GATING AT THE RIM OF THE GORGE Mutagenesis studies reveal that three residues, $W_{286(279)}$, $Y_{72(70)}$ and $Y_{124(121)}$ are critical for dictating specificity of BW284c51, decamethonium, and propidium (123) (Figure 5). These residues are also essential for binding of the peptide, fasciculin (Z Radić, R Duran, DC Vellom, Y Li, C Cervenansky & P Taylor, submitted). Decamethonium and BW284c51 likely span between the choline binding subsite and a portion of the peripheral anionic site whereas propidium and fasciculin are peripheral site selective ligands. In the case of BW284c51 a partitioning of free energy shows essentially linear free energy relationships for summing the contributions of the three residues

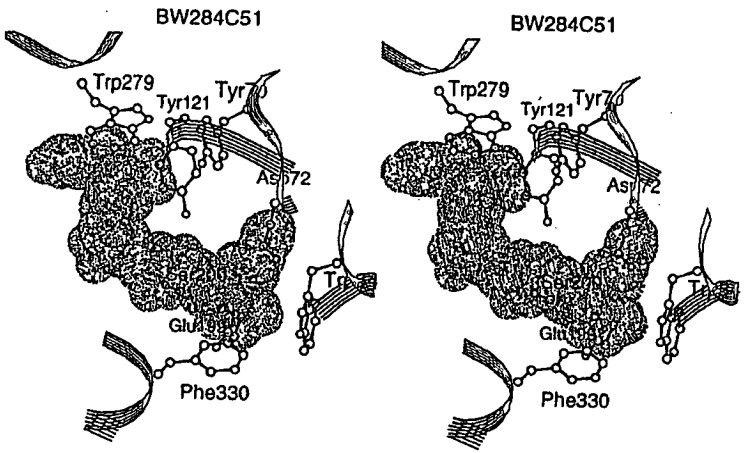


Figure 5 Positions of critical amino acid side chains for an energy minimized complex between BW284c51 and acetylcholinesterase (123).

to stabilization of the complex (123). W_{286} appears to be the most important residue for BW284c51, although each ligand shows slightly different partitioning of free energy. The involvement of only a small number of residues in stabilizing specific complexes is buttressed by the observation that a $BuChE_{1-174}$ $AChE_{175-575}$ chimera behaves like the $Y_{72}N,Y_{124}Q$ mutation and the $Y_{72}N,Y_{124}Q,W_{286}R$ mutation behaves similarly to mouse BuChE for inhibition by BW284c51. This is not the case for decamethonium and propidium, which suggests different binding loci for the latter ligands on the two respective cholinesterases (123).

$D_{74(72)}$ also affects the binding of these ligands (123, 125, 140) and it too is positioned rather close to the rim of the gorge. The site near the rim of the gorge defined by W_{286}, Y_{72}, Y_{124}, D_{74} has several features in common with the W_{86}, F_{337}, Y_{342}, E_{202} site found at the base of the gorge. Since *bis*-quaternary ligands span between the two sites, a similar complement of residues may thus be stabilizing each end of the *bis*-quaternary ligand.

$D_{74(72)}$ is conserved in BuChE as D_{70}. In fact, mutation of this residue to G is responsible for succinylcholine-induced paralysis in man (13, 14, 124, 125); an increased K_m and resistance to dibucaine inhibition and succinylcholine catalysis can be demonstrated in the mutant enzyme (140). Curiously, other mutations that concommitantly appear with the $D_{70}G$, H_{114}, Y_{561} and P_{425} restore some of the catalytic efficiency of the $D_{70}G$ mutation (140). Studies involving a BuChE template and replacement of residues with those found in AChE (135) and BuChE-AChE chimerae (141) have yielded results complementary to those obtained with the AChE template. In *Drosophila*, Y_{109} corresponds to D_{70} and mutations here influence inhibitor specificity (142).

Occupation of the peripheral site affects the conformation of the active center and the configuration of bound ligands at the active center (143, 144). Mutagenesis studies should further delineate the residues involved in this allosteric linkage (123, 138).

The three domains outlined above appear primarily responsible for the reported selectivity of AChE and BuChE for substrates and inhibitors. Specificity for acyl chain length and the propensity for substrate activation or inhibition are governed largely by the two phenylalanines, $F_{295(288)}$ and $F_{297(290)}$, whose side chains outline the acyl pocket. This region also governs the reactivity of isoOMPA for the enzyme; steric hindrance precludes isoOMPA from rapidly reacting with AChE. The BuChE selectivity of ethopropazine arises from its ability to be accommodated in the choline binding subsite. The diethylamino-2-propyl side chain exhibits interference with $F_{337(330)}$ in AChE whereas $A_{332(330)}$ in BuChE enables the fit (123). Finally, the site near the rim of the gorge dictates specificity of the *bis*-quaternary inhibitors and peptides that cannot fit at the base of the gorge; BW284c51, propidium, and fasciculin are the prime examples.

CARBOXYL-TERMINUS Several mutations of this region have emerged from a knowledge of the sequence and alternatively spliced forms. Gibney et al (126) documented the cassette characteristics of the individual exons. By splicing the invariant exons encoding the *Torpedo* enzyme to the exon encoding the glycophospholipid signal (exon 5) through loop-out mutagenesis, the glycophospholipid-linked form of AChE was synthesized in transfected COS cells. By dropping an intermediate exon, a truncated, but inactive, enzyme carrying the glycophospholipid-linkage was formed. By deleting the terminal exons (5 and 6), the expressed enzyme was secreted into the medium and lacked the glycophospholipid attachment. A natural splice variant in the mouse enzyme yields AChE with the same properties with virtually all of the enzyme appearing in the media (66, 67). Hence, the exon encoding the glycophospholipid linkage signal is both necessary and sufficient for generating the signal sequence for processing and addition of the glycophospholipid. Removal of the cysteine from exon 6 (128) or formation of a truncated hydrophilic form of the enzyme results in secretion of a monomeric enzyme (126). Similar dependencies of membrane attachments have been documented in *Drosophila* cholinesterase (145, 146).

An important development in the study of the assembly process has been the cloning of the gene that encodes the collagen-containing tail species in the *Torpedo* enzyme (147). Although there appear to be multiple tail subunits, coexpression of the cDNA encoding the catalytic subunit and that encoding the tail unit gave rise to the expected asymmetric species for both *Torpedo* and rat catalytic subunit cDNAs (28, 147). Moreover, truncation of the tail subunit cDNA showed that the amino-terminal portion of the tail molecule contains the sulfhydryl necessary for the intersubunit disulfide linkage (148).

Transfection of the cDNAs encoding the hydrophilic (exon 6) and glycophospholipid-linked (exon 5) forms of AChE generates the expected multiplicity of species seen in vivo (149). Hence, assembly to the various oligomeric species of AChE and processing occur with the transfected cDNAs. Transfection of mouse and human genomic constructs into various cell lines shows tissue selective splicing of mRNA to achieve a diversity of gene products (D Vellom, S Camp, and P Taylor, submitted).

GLYCOSYLATION Human AChE contains three N-linked glycosylation recognition sequences at N_{265}, N_{350}, and N_{464}. Deletion of the recognition sequences singly and in combination diminishes biosynthesis and secretion of the enzyme. The influence appears progressive since an enzyme deficient in all three signals shows the least expression, followed by the mutation with two of the three signals deleted (150). Glycosylation increases the thermal stability of the enzyme, but did not affect the catalytic parameters. Initial studies with the *Drosophila* enzyme also indicate that active enzyme can be synthesized in the absence of glycosylation (146, 151).

CHOLINESTERASE CHIMERAE Construction of cholinesterase chimerae has been useful in analyzing gene structure in relation to function and in identifying domains of the molecule responsible for particular functional characteristics. The initial approach deleted exon 5 and demonstrated secretion of the *Torpedo* enzyme (126); variants of this construct are discussed above. Attachment of the carboxyl-terminal signal sequence contained in exon 5 to upstream exons or to sequence encoding the amino-terminal portion of the collagen-containing tail unit yielded glycophospholipid-attached enzymes (126, 149). A second approach entailed forming BuChE chimerae between wild-type and naturally occurring mutants (125) and enabled examination of the influence of secondary mutations on the $D_{70}G$ mutation (125). Hence, portions of carboxyl-terminal domain of the molecule can modulate the consequences of an amino-terminal modification. Formation of active chimerae between AChE and BuChE have led to assigning domains responsible for inhibitor specificity, for delimiting the selection of residues in site-specific mutagenesis (131, 141). Comparisons of specificity between site-specific mutants and chimerae can often rule out an influence of several residues on inhibitor specificity.

Relationship of Ligand Binding Sites on Acetylcholinesterase to Those on Other Acetylcholine-Binding Proteins

Examination of the high resolution structure of AChE in relation to its functional characteristics and specificity of ligand binding sites may provide insights into the structure of other Ach binding proteins. We have already alluded to the similarities in both the proximal aromatic clusters and the more distant negative charges residing at the choline binding subsite and at the peripheral anionic site in AChE. Further parallels can be drawn with the aromatic clusters in the phosphorylcholine binding antibody and in chemically synthesized host ligands that bind quaternary ligands (87, 88, 152). Chemical labeling studies also show proximity of tyrosines and tryptophan in the vicinity of the ligand binding site on the acetylcholine receptor (153–155). Moreover, tacrine, a ligand that inhibits AChE by binding at the choline subsite (83, 123) also shows a propensity to inhibit K^+-channels. Mutagenesis studies are beginning to define the nature of a quaternary ammonium binding site within the K^+ channel, and a tyrosine substitution for threonine enhances tetraethylammonium inhibition of K^+ conductance (156). However, apart from using proximal aromatic residues and longer range electrostatic forces to stabilize the quaternary ligands and perhaps a more global organization of charges to form a macromolecular dipole to direct the binding of the ligand, there may be few specific parallels between the recognition sites on acetylcholine binding proteins.

In the case of the nicotinic acetylcholine receptor, the ligand binding site appears not to be in the central ion cavity or "gorge"; rather, agonists bind

at distinct sites at the periphery of the receptor (157). Entry of Ach to its two binding sites on the receptor appears to be normal to the axis defined by the ion permeability channel through the membrane. Finally, Ach binding sites are formed at subunit interfaces on the nicotinic receptor rather than being central to one of the subunits.

The muscarinic receptor presents an even different situation since the binding site must be constructed from within the seven membrane-spanning regions (158), a constraint not found for a globular protein or an extracellular domain of a membrane-associated protein.

Ach binds with relatively low affinity to an activatible state of the nicotinic receptor ($K_D \approx 10^{-4}$M), but short-term exposure results in desensitization and concomitant formation of a high-affinity state for Ach, $K_D = 5 \times 10^{-8}$M (153, 159, 160). This low dissociation constant may be contrasted with an AChE K_m of 0.5–1.0 $\times$ 10^{-4}M. Deconstruction of the AChE K_m would indicate that Ach dissociation constant (k_{-1}/k_1 in Scheme 1) is actually larger than K_m. Each state of the nicotinic receptor is designed to recognize the parent ligand whose acetoxy group is planar or trigonal, while in the case of AChE, the site is designed to force the formation of a transition state that is best approximated by a tetrahedral conformation around the carbonyl-containing carbon. The dissociation constant of the enzyme for this transition state, K_{TS}, can be estimated from $K_{TS} = K_m{}^{\circ}k_u/k_c$, where k_c/k_u (the ratio of catalyzed and uncatalyzed ester hydrolysis) is the catalytic enhancement provided by the enzyme. The product of K_m ($\sim 10^{-4}$ M) and k_u/k_c ($\sim 10^{-13}$) (113) yields a value of $\sim 10^{-17}$M and reflects a uniquely high affinity for the labile transition state of the substrate. Hence, receptors and AChE are designed to recognize and catalytically force or accommodate distinct conformations of acetylcholine. Accordingly, these unique binding characteristics are likely to be reflected in major differences in molecular and spatial characteristics of their respective binding sites.

Literature Cited

1. Clermont P. 1854. Note sur la préparation de quelques éthers. *CR Acad. Sci. Paris* 39:338–41
2. Argyll-Robinson D. 1863. The calibar bean as a new agent in ophthalmic practice. *Edinburgh Med. J.* 8:815–20
3. Dale HH. 1914. The action of certain esters of choline and their relation to muscarine. *J. Pharmacol. Exp. Ther.* 6:147–90
4. Loewi O, Navratil E. 1926. Über humorale Übertragbarkeit der Herznervenwirkung. XI. Über der Mechanismus der Vaguswirkung von Physostigmin und Ergotamin. *Pfluegers Arch.* 214:689–96
5. Ellman GL, Courtney KD, Andres V Jr, Featherstone RM. 1961. A new and rapid colorimetric determination of acetylcholinesterase activity. *Biochem. Pharmacol.* 7:88–95
6. Koelle GB, Friedenwald JS. 1949. A histochemical method for localizing cholinesterase activity. *Proc. Soc. Exp. Biol. Med.* 70:617–22
7. Karnovsky MS, Roots L. 1964. A direct coloring thiocholine method for cholinesterase. *J. Histochem. Cytochem.* 12:219–32
8. Taylor P. 1990. Anticholinesterase agents. In *The Pharmacological Basis of Therapeutics*, ed. AG Gilman, LS Goodman, TW Rall, F Murad, AS Nies, P Taylor, pp. 131–49. New York: Pergamon. 8th ed.
9. Schumacher M, Camp S, Maulet Y, Newton M, MacPhee-Quigley K, et al. 1986. Primary structure of *Torpedo californica* acetylcholinesterase deduced from cDNA sequences. *Nature* 319:407–9
10. Sussman JL, Harel M, Frolow F, Oefner C, Goldman A, et al. 1991. Atomic structure of acetylcholinesterase from *Torpedo californica*: a prototypic acetylcholine-binding protein. *Science* 253:872–79
11. Quinn DM. 1987. Acetylcholinesterase: enzyme structure, reaction dynamics, and virtual transition states. *Chem. Rev.* 87:955–79
12. Rosenberry TL. 1975. Acetylcholinesterase. *Adv. Enzymol. Relat. Areas Mol. Biol.* 43:103–218
13. Whittaker M. 1986. Cholinesterases. In *Monographs in Human Genetics*, ed. RS Sparkes, 11:125. Basel: Karger
14. Soreq H, Zakut H. 1990. Cholinesterase genes: multileveled regulation. In *Monographs in Human Genetics*, ed. RS Sparkes, 13:1–102. Basel: Karger
15. Sussman JL, Silman I. 1992. Acetylcholinesterase: structure and use as a model for specific cation protein interactions. *Curr. Opin. Struct. Biol.* 2:721–29
16. Taylor P. 1991. The cholinesterases. *J. Biol. Chem.* 266:4025–28
17. Massoulié J, Pezzementi L, Bon S, Krejci E, Vallette FM. 1993. Molecular and cellular biology of the cholinesterases. *Prog. Neurobiol.* 41: 31–91
18. Shafferman A, Velan B, eds. 1992. *Multidisciplinary Approaches to Cholinesterase Functions*. New York: Plenum. 293 pp.
19. Swillens S, Ludgate M, Mercken L, Dumont JE, Vassart G. 1986. Analysis of sequence and structure homologies between thyroglobulin and acetylcholinesterase: possible functional and clinical significance. *Biochem. Biophys. Res. Commun.* 137:142–48
20. Hall LMC, Spierer P. 1986. The *Ace* locus of *Drosophila melanogaster*: structural gene for acetylcholinesterase with an unusual 5′ leader. *EMBO J.* 5:2949–54
21. Lockridge O, Bartels CF, Vaughan TA, Wong CK, Norton SE, Johnson LL. 1987. Complete amino acid sequence of human serum cholinesterase. *J. Biol. Chem.* 262:549–57
22. Prody CA, Zevin-Sonkin D, Gnatt A, Goldberg O, Soreq H. 1987. Isolation and characterization of full-length cDNA clones coding for cholinesterase from fetal human tissues. *Proc. Natl. Acad. Sci. USA* 84:3555–59
23. McTiernan C, Adkins S, Chatonnet A, Vaughan TA, Bartels CF, et al. 1987. Brain cDNA clone for human cholinesterase. *Proc. Natl. Acad. Sci. USA* 84:6682–86
24. Rachinsky TL, Camp S, Li Y, Ekström TJ, Newton M, Taylor P. 1990. Molecular cloning of mouse acetylcholinesterase: tissue distribution of alternatively spliced mRNA species. *Neuron* 5:317–27
25. Doctor BP, Chapman TC, Christner CE, Deal CD, De La Hoz DM, et al. 1990. Complete amino acid sequence of fetal bovine serum acetylcholinesterase and its comparison in various regions with other cholinesterases. *FEBS Lett.* 266:123–27
26. Soreq H, Ben-Aziz R, Prody CA, Seidman S, Gnatt A, et al. 1990. Molecular cloning and construction of the coding region for human acetylcholinesterase reveals a G+C-rich attenuating structure. *Proc. Natl. Acad. Sci. USA* 87:9688–92
27. Chatonnet A, Lorca T, Barakat A, Aron E, Jbilo O. 1991. Structure of rabbit butyrylcholinesterase gene deduced from genomic clones and from a cDNA with introns. *Cell Mol. Neurobiol.* 11:119–30
28. Legay C, Bon S, Vernier P, Coussen F, Massoulié J. 1993. Cloning and expression of a rat acetylcholinesterase subunit: generation of multiple molecular forms and complementarity with a *Torpedo* collagenic subunit. *J. Neurochem.* 60:337–46
29. Hall LMC, Malcolm CA. 1991. The acetylcholinesterase gene of *Anopheles stephensi*. *Cell. Mol. Neurobiol.* 11: 131–41
30. Maulet Y, Ballivet M. 1991. Two

different genes encoding cholinesterases in chicken. See Ref. 162, p. 186
31. Rubino S, Mann SKO, Hori RT, Pinko C, Firtel RA. 1989. Molecular analysis of a developmentally regulated gene required for *Dictyostelium* aggregation. *Dev. Biol.* 131:27–36
32. Bomblies L, Biegelmann E, Döring V, Gerisch G, Krafft-Czepa H, et al. 1990. Membrane-enclosed crystals in *Dictyostelium discoideum* cells, consisting of developmentally regulated proteins with sequence similarities to known esterases. *J. Cell Biol.* 110:669–79
33. Oakeshott JG, Collet C, Phillis RW, Nielsen KM, Russell RJ, et al. 1987. Molecular cloning and characterization of esterase-6, a serine hydrolase of *Drosophila*. *Proc. Natl. Acad. Sci. USA* 84:3359–63
34. Collet C, Nielsen KM, Russell RJ, Karl M, Oakeshott JG, Richmond RC. 1990. Molecular analysis of duplicated esterase genes in *Drosophila melanogaster*. *Mol. Biol. Evol.* 7:9–28
35. Mouches C, Pauplin Y, Agarwal M, Lemieux L, Herzog M, et al. 1990. Characterization of amplification core and esterase B1 gene responsible for insecticide resistance in *Culex*. *Proc. Natl. Acad. Sci. USA* 87:2574–78
36. Hanzlik TN, Abdel-Aal YAI, Harshman LG, Hammock BD. 1989. Isolation and sequencing of cDNA clones coding for juvenile hormone esterase from *Heliothis virescens*. Evidence for a catalytic mechanism for the serine carboxylesterases different from that of the serine proteases. *J. Biol. Chem.* 264:12419–25
37. Shimada Y, Sugihara A, Tominaga Y, Izumi T, Tsunasawa S. 1989. cDNA molecular cloning of *Geotrichum candidum* lipase. *J. Biochem.* 106:383–88
38. Long RM, Satoh H, Martin BM, Kimura S, Gonzalez FJ, Pohl LR. 1988. Rat liver carboxylesterase: cDNA cloning, sequencing, and evidence for a multigene family. *Biochem. Biophys. Res. Commun.* 156:866–73
39. Korza G, Ozols J. 1988. Complete covalent structure 60-kDa esterase isolated from 2,3,7,8-tetrachlorodibenzo-p-dioxin-induced rabbit liver microsomes. *J. Biol. Chem.* 263:3486–95
40. Han JH, Stratowa C, Rutter WJ. 1987. Isolation of full-length putative rat lysophospholipase cDNA using improved methods for mRNA isolation and cDNA cloning. *Biochemistry* 26: 1617–25
41. Kyger EM, Wiegand RC, Lange LG. 1989. Cloning of the bovine pancreatic cholesterol esterase/lysophospholipase. *Biochem. Biophys. Res. Commun.* 164: 1302–9
42. Ollis DL, Cheah E, Cygler M, Dijkstra B, Frolow F, et al. 1992. The α/β hydrolase fold. *Prot. Eng.* 5:197–211
43. Cooper A, Russey H. 1989. Characterization of the yeast KEX1 gene product: a carboxypeptidase involved in processing secreted precursor proteins. *Mol. Cell. Biol.* 9:2706–14
44. Franken SM, Rozeboom HJ, Kalk KH, Dijkstra BW. 1991. Crystal structure of haloalkane dehalogenase: an enzyme to detoxify halogenated alkanes. *EMBO J.* 10:1297–302
45. Pathak D, Ngai KL, Ollis D. 1988. X-ray crystallographic structure of dienelactone hydrolase at 2.8 Å. *J. Mol. Biol.* 204:435–45
46. Olson PF, Fessler LI, Nelson RE, Campbell AG, Fessler JH. 1990. Glutactin, a novel *Drosophila* basement membrane related glycoprotein with sequence similarity to serine esterases. *EMBO J.* 9:3593–601
47. De la Escalera S, Backamp E-O, Moya F, Piovant M, Jimenez F. 1990. Characterization and gene cloning of neurotactin, a *Drosophila* transmembrane protein related to cholinesterases. *EMBO J.* 9:3593–601
48. Gentry MK, Doctor BP. 1991. Alignment of amino acid sequences of acetylcholinesterases and butyrylcholinesterases. See Ref. 162, pp. 394–98
49. Cygler M, Schrag J, Sussman JL, Harel M, Silman I, et al. 1993. Relationship between sequence conservation and three-dimensional structure in a large family of esterases, lipases, and related proteins. *Protein Sci.* 2: 366–82
50. MacPhee-Quigley K, Vedvick TS, Taylor P, Taylor SS. 1986. Profile of disulfide bonds in acetylcholinesterase. *J. Biol. Chem.* 261:13565–70
51. Lockridge O, Adkins S, La Du BN. 1987. Location of disulfide bonds within the sequence of human serum cholinesterase. *J. Biol. Chem.* 262: 12945–52
52. MacPhee-Quigley K, Taylor P, Taylor SS. 1985. Primary structures of the catalytic subunits from two molecular forms of acetylcholinesterase: a comparison of NH2-terminal and active center sequences. *J. Biol. Chem.* 260: 12185–89
53. Gibney G, Camp S, Dionne M, MacPhee-Quigley K, Taylor P. 1990. Mutagenesis of essential functional residues in acetylcholinesterase. *Proc. Natl. Acad. Sci. USA* 87:7546–50
54. Augustinsson KB. 1948. Cholinesterases: a study in comparative enzymology. *Acta Physiol. Scand.* 15 (Suppl. 52):1–182
55. Silver A. 1974. *The Biology of Cholinesterases*. Amsterdam: North-Holland
56. Gnagey AL, Forte M, Rosenberry TL. 1987. Isolation and characterization of acetylcholinesterase from *Drosophila*. *J. Biol. Chem.* 262:1140–45
57. Toutant J-P, Massoulié J, Bon S. 1985. Polymorphism of pseudocholinesterase in *Torpedo marmorata* tissues: comparative study of the catalytic and molecular properties of this enzyme with acetylcholinesterase. *J. Neurochem.* 44:580–92
58. Johnson CD, Rand JB, Herman RK, Stern BD, Russell RL. 1988. The acetylcholinesterase genes of *C. elegans*: identification of a third gene (ace-3) and mosaic mapping of a synthetic lethal phenotype. *Neuron* 1:165–73
59. Arpagaus M, Richier P, L'Hermitte Y, Le Roy F, Bergé J, et al. 1992. Nematode acetylcholinesterases: several genes and molecular forms of their products. See Ref. 18, pp. 65–74
60. Maulet Y, Camp S, Gibney G, Rachinsky TL, Ekström TJ, Taylor P. 1990. A single gene encodes glycophospholipid-anchored and asymmetric acetylcholinesterase forms: alternative coding exons contain inverted repeat sequences. *Neuron* 4:289–301
61. Li Y, Camp S, Rachinsky TL, Getman DK, Taylor P. 1991. Gene structure of mammalian acetylcholinesterase: alternative exons dictate tissue-specific expression. *J. Biol. Chem.* 266:23083–90
62. Arpagaus M, Kott M, Vatsis KP, Bartels CF, La Du BN, Lockridge O. 1990. Structure of the gene for human butyrylcholinesterase. Evidence for a single copy. *Biochemistry* 29:124–31
63. Sikorav JL, Krejci E, Massoulié J. 1987. cDNA sequences of *Torpedo marmorata* acetylcholinesterase: primary structure of the precursor of a catalytic subunit; existence of multiple 5'-untranslated regions. *EMBO J.* 6: 1865–73
64. Sikorav JL, Duval N, Anselmet A, Bon S, Krejci E, et al. 1988. Complex alternative splicing of acetylcholinesterase transcript in *Torpedo* electric organ; primary structure of the precursor of the glycolipid-anchored dimeric form. *EMBO J.* 7:2983–93
65. Schumacher M, Maulet Y, Camp S, Taylor P. 1988. Multiple messenger RNA species give rise to the structural diversity in acetylcholinesterase. *J. Biol. Chem.* 263:18979–87
66. Li Y, Camp S, Taylor P. 1993. Tissue-specific expression and alternative mRNA processing of the mammalian acetylcholinesterase gene. *J. Biol. Chem.* 268:5790–97
67. Li Y, Camp S, Rachinsky TL, Bongiorno C, Taylor P. 1993. Promoter elements and transcriptional control of the mouse acetylcholinesterase gene. *J. Biol. Chem.* 268:3563–72
68. Gibney G, MacPhee-Quigley K, Thompson B, Vedvick T, Low MG, et al. 1988. Divergence in primary structure between molecular forms of acetylcholinesterase. *J. Biol. Chem.* 263:1140–45
69. Rotundo RL, Gomez AM, Fernandez-Valle C, Randall WR. 1988. Allelic variants of acetylcholinesterase forms in avian nerves and muscle and encoded by a single gene. *Proc. Natl. Acad. Sci. USA* 52:7121–25
70. Getman DK, Eubanks J, Evans G, Taylor P. 1992. Assignment of the human acetylcholinesterase gene to chromosome 7q22. *Am. J. Hum. Genet.* 50:170–77
71. Ehrlich G, Viegas-Pequignot E, Ginzberg D, Sindel L, Soreq H, Zakut H. 1992. Mapping the human acetylcholinesterase gene to chromosome 7q22 by fluorescent in situ hybridization coupled with selective PCR amplification from a somatic hybrid cell panel and chromosome-sorted DNA libraries. *Genomics* 13:1192–97
72. Gaughan G, Park H, Priddle J, Craig I, Craig S. 1991. Refinement of the localization of human butyrylcholinesterase to chromosome 3q26.1-q26.2 using a PCR-derived probe. *Genomics* 11:455–58
73. Allderdice PW, Gardner HAR, Galutira D, Lockridge O, La Du BN, McAlpine PJ. 1991. The cloned butyrylcholinesterase (*BCHE*) gene maps to a single chromosome site, 3q26. *Genomics* 11:452–54
74. Rachinsky TL, Crenshaw EB III, Taylor P. 1992. Assignment of the gene for acetylcholinesterase to distal mouse-chromosome 5. *Genomics* 14:511–14
75. Sussman JL, Harel M, Frolow F, Varon L, Toker L, et al. 1988. Purification and crystallization of a dimeric form of acetylcholinesterase from *Torpedo*

californica subsequent to solubilization with phosphatidylinositol-specific phospholipase. *J. Mol. Biol.* 203:821–23
76. Schrag JD, Schmid MF, Morgan DG, Phillips GN, Wah C, Tang L. 1988. Crystallization and preliminary X-ray diffraction analysis of 11S acetylcholinesterase. *J. Biol. Chem.* 263: 9795–801
77. Schrag JD, Li Y, Wu S, Cygler M. 1991. Ser-His-Glu triad forms the catalytic site of the lipase from *Geotrichum candidum*. *Nature* 351: 761–64
78. Harel M, Sussman JL, Krejci E, Bon S, Chanal P, et al. 1992. Conversion of acetylcholinesterase to butyrylcholinesterase, modeling and mutagenesis. *Proc. Natl. Acad. Sci. USA* 89: 10827–31
79. Rosenberry TL, Neumann E. 1977. Interaction of ligands with acetylcholinesterase. Use of temperature-jump relaxation kinetics in the binding of specific fluorescent ligands. *Biochemistry* 16:3870–78
80. Nolte HJ, Rosenberry TL, Neumann E. 1980. Effective charge on acetylcholinesterase active sites determined from the ionic strength dependence of association rate constants with cationic ligands. *Biochemistry* 19:3705–11
81. Tan RC, Truong TN, McCammon JA, Sussman JL. 1993. Acetylcholinesterase: electrostatic steering increases the rate of ligand binding. *Biochemistry* 32:401–3
82. Ripoll DR, Faerman CH, Axelsen PH, Silman I, Sussman JL. 1993. An electrostatic mechanism for substrate guidance down the aromatic gorge of acetylcholinesterase. *Proc. Natl. Acad. Sci. USA* 90:5128–32
83. Sussman JL, Harel M, Silman I. 1992. Three dimensional structure of acetylcholinesterase. See Ref. 18, pp. 95–108
84. Harel M, Schalk I, Ehret-Sabatier L, Bouet F, Goeldner M, et al. 1993. Quaternary ligand binding site of acetylcholinesterase as revealed by X-ray crystallography and photoaffinity labelling. *Proc. Natl. Acad. Sci. USA.* 90:9031–35
85. Hasan FB, Cohen SG, Cohen JB. 1980. Hydrolysis by acetylcholinesterase. Apparent molal volumes and trimethyl and methyl subsites. *J. Biol. Chem.* 255:3898–904
86. Cohen SG, Salih E, Solomon M, Howard S, Chishti SB, Cohen JB. 1989. Reactions of 1-bromo-2-[^{14}C]pinacolone with acetylcholinesterase from *Torpedo nobiliana*. Effects of 5-trimethylammonio-2-pentanone and diisopropyl fluorophosphate. *Biochem. Biophys. Acta* 997:167–75
87. Dougherty DA, Stauffer DA. 1990. Acetylcholine binding by a synthetic receptor: implication for biological recognition. *Science* 250:1558–60
88. Segal DM, Padlan EA, Cohen GH, Rudikoff S, Potter M, Davies DR. 1974. The three-dimensional structure of phosphorylcholine—binding mouse immunoglobulin Fab and the nature of the antigen binding site. *Proc. Natl. Acad. Sci. USA* 71:4295–302
89. Pressman D, Grossberg AL, Pence LH, Pauling L. 1946. The reaction of antiserum homologous to the p-azo phenyltrimethylammonium group. *J. Am. Chem. Soc.* 68:250–55
90. Changeux J-P. 1966. Responses of acetylcholinesterase from *Torpedo marmorata* to salts and curarizing drugs. *Mol. Pharmacol.* 2:369–92
91. Taylor P, Lappi S. 1975. Interaction of fluorescence probes with acetylcholinesterase. The site and specificity of propidium binding. *Biochemistry* 14: 1989–97
92. Berman HA, Yguerabide J, Taylor P. 1980. Fluorescence energy transfer on acetylcholinesterase: spatial relationship between peripheral site and active center. *Biochemistry* 19:2226–35
93. Weise C, Kreienkamp HJ, Raba R, Pedak A, Aaviksaar A, Hucho F. 1990. Anionic subsites of the acetylcholinesterase from *Torpedo californica:* affinity labelling with the cationic reagent *N,N*-dimethyl-2-phenyl-aziridinium. *EMBO J.* 9:3885–88
94. Amitai G, Taylor P. 1991. Characterization of peripheral anionic site peptides of AChE by photoaffinity labeling with monoazidopropidium (MAP). See Ref. 162, p. 285
95. Haas R, Adams EW, Rosenberry MA, Rosenberry TL. 1992. Substrate-selective inhibition and peripheral site labeling of acetylcholinesterase by platinum(terpyridine)chloride. See Ref. 18, pp. 131–40
96. Marchot P, Khelif A, Ji YH, Mansuelle P, Bougis PE. 1993. Binding of ^{125}I-fasciculin to rat brain acetylcholinesterase. The complex still binds diisopropyl fluorophosphate. *J. Biol. Chem.* 268:12458–67
97. Cervenansky C, Dajas F, Harvey AL, Karlsson E. 1991. The fasciculins. In *Snake Toxins*, ed. AC Harvey, pp. 303–21. New York: Pergamon
98. Radić Z, Reiner E, Taylor P. 1991. Role of the peripheral anionic site on acetylcholinesterase: inhibition by substrates and coumarin derivatives. *Mol. Pharmacol.* 39:98–104
99. Berman HA, Nowak MW. 1992. Influence of ion composition of the medium on acetylcholinesterase conformation. See Ref. 18, pp. 149–56
100. Wilson IB, Quan C. 1958. Acetylcholinesterase studies on molecular complementariness. *Arch. Biochem. Biophys.* 73:131–38
101. Kreienkamp HJ, Weise C, Raba R, Aaviksaar A, Hucho F. 1991. Anionic subsites of the catalytic center of acetylcholinesterase from *Torpedo* and from cobra venom. *Proc. Natl. Acad. Sci. USA* 88:6117–21
102. Schalk I, Ehret-Sabatier L, Bouet F, Goeldner M, Hirth C. 1992. Structural analysis of acetylcholinesterase ammonium binding sites. See Ref. 18, pp. 117–20
103. Shinitzky M, Dudai Y, Silman I. 1973. Spectral evidence for the presence of tryptophan in the binding site of acetylcholinesterase. *FEBS Lett.* 30: 125–28
104. Mooser G, Sigman DS. 1974. Ligand binding properties of acetylcholinesterase determined with fluorescent probes. *Biochemistry* 13:2299–307
105. Barnett P, Rosenberry TL. 1977. Catalysis by acetylcholinesterase: acceleration of the hydrolysis of neutral acetic acid esters by certain aromatic cations. *J. Biol. Chem.* 252:7200–6
106. Berman HA, Leonard K. 1989. Chiral reactions of acetylcholinesterase probed with enantiomeric methylphosphonothioates. Noncovalent determinants of enzyme chirality. *J. Biol. Chem.* 264:3942–50
107. Belleau B, DiTullio V. 1970. Kinetic effects of alkyl quaternary ammonium salts on the methanesulfonylation of the acetylcholinesterase catalytic center. Significance of substituent volumes and binding enthalpies. *J. Am. Chem. Soc.* 92:6320–25
108. Belleau B, DiTullio V, Tsai Y-H. 1970. Kinetic effects of leptocurares and pachycurares on the methanesulfonylation of acetylcholinesterase. A correlation with pharmacodynamic properties. *Mol. Pharmacol.* 6:41–45
109. Wilson IB. 1967. Conformation changes in acetylcholinesterase. *Ann. NY Acad. Sci.* 144:664–79
110. Taylor P, Jacobs NM. 1974. Interaction between bisquaternary ammonium ligands and acetylcholinesterase: complex formation studied by fluorescence quenching. *Mol. Pharmacol.* 10:93–107
111. Wee VT, Sinha BK, Taylor P, Chignell CF. 1976. Interaction of spin-labeled bisquaternary ammonium ligands with acetylcholinesterase. *Mol. Pharmacol.* 12:667–77
112. Aldridge WN, Reiner E. 1972. *Enzyme Inhibitors as Substrates*. Amsterdam: Elsevier. 328 pp.
112a. Himel CM, Taylor JL, Pape C, Millar DB, Christopher J, Kurlansik L. 1979. Acridine araphanes: a new class of probe materials for biological receptors. *Science* 205:1277–79
113. Bazclyansky M, Robey E, Kirsch JF. 1986. Fractional diffusion limited component of reactions catalyzed by acetylcholinesterase. *Biochemistry* 25: 125–32
114. Quinn DM, Pryor AN, Selwood T, Lee BH, Acheson SA, Barlow PN. 1991. The chemical mechanism of acetylcholinesterase reactions. Biological catalysis at the speed limit. See Ref. 162, pp. 252–57
115. Froede HC, Wilson IB. 1984. Direct determination of acetyl-enzyme intermediate in the acetylcholinesterase catalysis of acetylcholine and acetylthiocholine. *J. Biol. Chem.* 259:11010–13
116. Quinn DM, Selwood T, Pryor AN, Lee BH, Leu L-S, et al. 1992. Cryptic catalysis and cholinesterase function. See Ref. 18, pp. 141–48
117. Albery WJ, Knowles JR. 1976. Evolution of enzyme function and the development of catalytic efficiency. *Biochemistry* 15:5631–40
118. Rosenberry TL. 1975. Catalysis by acetylcholinesterase: Evidence that the rate-limiting step for acylation with certain substrates precedes general acid-base catalysis. *Proc. Natl. Acad. Sci. USA* 72:3834–38
119. Pryor AN, Selwood T, Leu L-S, Andracki MA, Lee BH, et al. 1992. Simple general acid-base catalysis of physiological acetylcholinesterase reactions. *J. Am. Chem. Soc.* 114:3896–900
120. Krupka RM, Laidler KJ. 1961. Molecular mechanisms for hydrolytic enzyme action II. Inhibition of acetylcholinesterase by excess substrate. *J. Am. Chem. Soc.* 83:1448–54
121. Augustinsson KB, Bartfai T, Mannervik B. 1974. A steady state model for butyrylcholinesterase from horse plasma. *Biochem. J.* 141:825–34
122. Cauet G, Friboulet A, Thomas D. 1987. Horse serum butyrylcho-

linesterase kinetics: a molecular mechanism based on inhibition studies with dansylaminoethyl trimethylammonium. *Biochem. Cell. Biol.* 65:529–35
123. Radic Z, Pickering N, Vellom DC, Camp S, Taylor P. 1993. Three distinct domains in the cholinesterase molecule confer selectivity for acetylcholinesterase and butyrylcholinesterase inhibitors. *Biochemistry*. 32:12074–84
124. La Du BN, Bartels CF, Nogueira CP, Arpagaus M, Lockridge O. 1991. Proposed nomenclature for human butyrylcholinesterase genetic variants identified by DNA sequencing. *Cell. Mol. Neurobiol.* 11:79–90
125. Neville LF, Gnatt A, Padan R, Seidman S, Soreq H. 1990. Anionic site interactions in human butyrylcholinesterase disrupted by two single point mutations. *J. Biol. Chem.* 265: 20735–38
126. Gibney G, Taylor P. 1990. Biosynthesis of *Torpedo* acetylcholinesterase in mammalian cells. Functional expression and mutagenesis of the glycophospholipid-anchored form. *J. Biol. Chem.* 265:12576–83
127. Soreq H, Seidman S. 1992. *Xenopus* oocyte microinjection: from gene to protein. *Methods Enzymol.* 207:225–65
128. Velan B, Grosfeld H, Kronman C, Leitner M, Gozes Y. 1991. The effect of elimination of intersubunit disulfide bonds on the activity, assembly, and secretion of recombinant human acetylcholinesterase: expression of acetylcholinesterase CYS-580-Ala mutant. *J. Biol. Chem.* 266: 23977–84
129. Shafferman A, Kronman C, Flashner Y, Leitner M, Grosfeld H, et al. 1992. Mutagenesis of human acetylcholinesterase; identification of residues involved in catalytic activity and in polypeptide folding. *J. Biol. Chem.* 267:17640–48
130. Shafferman A, Velan B, Ordentlich A, Kronman C, Grosfeld H, et al. 1992. Substrate inhibition of acetylcholinesterase: residues involved in signal transduction from the surface to the catalytic center. *EMBO J.* 11: 3561–68
131. Vellom DC, Radic Z, Li Y, Pickering NA, Camp S, Taylor P. 1993. Amino acid residues controlling acetylcholinesterase and butyrylcholinesterase
binant acetylcholinesterase in a *Baculovirus* system: kinetic properties of glutamate 199 mutants. *Biochemistry* 31:9760–67
133. Fischer M, Ittah A, Liefer I, Gorecki M. 1993. Expression and recognition of biologically active human acetylcholinesterase from *E. coli*. *Cell Mol. Neurobiol.* 13:25–38
134. Duval N, Bon S, Silman I, Sussman JL, Massoulié J. 1992. Site-directed mutagenesis of active site-related residues in *Torpedo* acetylcholinesterase. Presence of a glutamic acid in the catalytic triad. *FEBS Lett.* 309:421–23
135. Gnatt A, Lowenstein Y, Yaron A, Schwarz M, Soreq H. 1993. Site-directed mutagenesis of active site residues reveals plasticity of human butyrylcholinesterase in substrate and inhibitor interactions. *J. Neurochem.* In press
136. Krejci E, Duval N, Chatonnet A, Vincens P, Massoulié J. 1991. Cholinesterase-like domains in enzymes and structural proteins: functional and evolutionary relationships and identification of a catalytically essential aspartic acid. *Proc. Natl. Acad. Sci. USA* 88:6647–51
137. Chatel JM, Grassi J, Frobert Y, Massoulié J, Vallette FM. 1993. Existence of an inactive pool of acetylcholinesterase in chicken brain. *Proc. Natl. Acad. Sci. USA* 90:2476–80
138. Ordentlich A, Barak D, Kronman C, Flashner Y, Leitner M, et al. 1993. Dissection of the human acetylcholinesterase active center—Determinants of substrate specificity. *J. Biol. Chem.* 268:17083–95
139. Fournier D, Bride J-M, Hoffmann F, Karch F. 1992. Acetylcholinesterase: two types of modifications confer resistance to insecticide. *J. Biol. Chem.* 267:14270–74
140. Neville LF, Gnatt A, Loewenstein Y, Seidman S, Ehrlich G, Soreq H. 1992. Intramolecular relationships in cholinesterases revealed by oocyte expression of site-directed and natural variants of human BCHE. *EMBO J.* 11:1641–49
141. Lowenstein Y, Gnatt A, Neville LF, Soreq H. 1993. A chimeric human cholinesterase: identification of interaction sites responsible for recognition of acetyl-, butyrylcholin-

of cholinesterases: importance of tyrosine 109 in *Drosophila* protein. *NeuroReport* 3:39–42
143. Epstein DJ, Berman HA, Taylor P. 1979. Ligand-induced conformational changes in acetylcholinesterase investigated with fluorescent phosphonates. *Biochemistry* 18:4749–54
144. Berman HA, Becktel W, Taylor P. 1981. Spectroscopic studies on acetylcholinesterase: influence of peripheral-site occupation on active-center conformation. *Biochemistry* 20: 4803–10
145. Fournier D, Mutero A, Rungger D. 1992. *Drosophila* acetylcholinesterase: expression of a functional precursor in *Xenopus* oocytes. *Eur. J. Biochem.* 203:513–19
146. Fournier D, Mutero A, Pravalorio M, Bride J-M. 1992. *Drosophila* acetylcholinesterase: analysis of structure and sensitivity to insecticides by in vitro mutagenesis and expression. See Ref. 18, pp. 75–82
147. Krejci E, Coussen F, Duval N, Chatel JM, Legay C, et al. 1991. Primary structure of a collagenic tail subunit of *Torpedo* acetylcholinesterase: co-expression with catalytic subunit induces the production of collagen-tailed forms in transfected cells. *EMBO J.* 10:1285–93
148. Duval N, Krejci E, Grassi J, Coussen F, Massoulié J, Bon S. 1992. Molecular architecture of acetylcholinesterase collagen-tailed forms: construction of a glycolipid-tailed tetramer. *EMBO J.* 9:3255–61
149. Duval N, Massoulié J, Bon S. 1992. H and T subunits of acetylcholinesterase from *Torpedo*, expressed in COS cells, generate all types of globular forms. *J. Cell Biol.* 118:641–53
150. Velan B, Kronman C, Ordentlich A, Flashner Y, Leitner M, et al. 1993. N-glycosylation of human acetylcholinesterase effects enzyme stability and secretion efficiency but not enzymatic activity. *Biochem. J.* In press
151. Mutero A, Fournier D. 1992. Post-translational modifications of *Drosophila* acetylcholinesterase. In vitro mutagenesis and expression in *Xenopus* oocytes. *J. Biol. Chem.* 267: 1695–700
152. McCury A, Jimenez L, Stauffer DA, Dougherty DA. 1992. Biomimetic catalysis of S_N2 reactions through cation-π interactions. The role of polarizability in catalysis. *J. Am. Chem. Soc.* 114:10314–21
153. Changeux J-P, Galzi J-L, Devillers-Thiery A, Bertrand D. 1992. The functional architecture of the acetylcholine nicotinic receptor explored by affinity labelling and site-directed mutagenesis. *Q. Rev. Biophys.* 25:395–432
154. Abramson SN, Li Y, Culver P, Taylor P. 1989. An analog of lophotoxin reacts covalently with Tyr^{190} in the α-subunit of the nicotinic acetylcholine receptor. *J. Biol. Chem.* 264:12666–72
155. Cohen JB, Sharp SD, Liu WS. 1991. Structure of the agonist-binding site of the nicotinic acetylcholine receptor: [^{3}H]acetylcholine mustard identifies residues in the cation-binding subsite. *J. Biol. Chem.* 266:23354–64
156. Heginbotham L, MacKinnon R. 1992. The aromatic binding site for tetraethylammonium ion on potassium channels. *Neuron* 8:483–91
157. Unwin NJM. 1993. Nicotinic acetylcholine receptor at 9 Å resolution. *J. Mol. Biol.* 229:1101–24
158. Hulme EC, Kurtenbach E, Curtis CA. 1991. Muscarinic acetylcholine receptors: structure and function. *Biochem. Soc. Trans.* 19:133–38
159. Taylor P, Johnson DA, Brown RD. 1983. The linkage between ligand occupation and response of the nicotinic acetylcholine receptor. In *Advances in Membranes and Transport*, ed. A Kleinzeller, BR Martin, pp. 407–43. New York: Academic
160. Sine SM, Claudio T, Sigworth FJ. 1990. Activation of *Torpedo* acetylcholine receptors in mouse fibroblasts: Single channel current kinetics reveal distinct agonist binding affinities. *J. Gen. Physiol.* 153:305–12
161. Bartels CF, Zelinski T, Lockridge O. 1993. Mutation at codon 322 in the human acetylcholinesterase (ACHE) gene accounts for YT blood group polymorphism. *Am. J. Hum. Genet.* 52:928–36
162. Massoulié J, Bacou F, Barnard E, Chatonnet A, Doctor BP, Quinn DM, eds. 1991. *Cholinesterases: Structure, Function, Mechanism, Genetics and Cell Biology*. Washington, DC: Am. Chem. Soc. 414 pp.
163. Mason P, Adkins S, Gouet P, Lockridge O. 1993. Recombinant human butyrylcholinesterase G39OV, the fluoride-2 variant expressed in Chinese hamster ovary cells, is a low affinity variant. *J. Biol. Chem.* 268:14329–41
164. Fournier D, Mutero A, Pravalorio M,

Bride JM. 1993. *Drosophila* acetylcholinesterase: mechanism of resistance to organophosphates. *Chem. Biol. Interact.* 87:233–38

165. Kerem A, Kronman C, Bar-Nun S, Shafferman A, Velan B. 1993. Interrelations between assembly and secretion of recombinant human acetylcholinesterase. *J. Biol. Chem.* 268: 180–84

166. Saxena A, Doctor BP, Maxwell DM, Lenz DE, Radicae Z, Taylor P. 1993. The role of glutamate 199 in the aging of cholinesterase. *Biochem. Biophys. Res. Commun.* In Press

Annu. Rev. Pharmacol. Toxicol. 1994. 34:321–41

ABSORPTION OF PEPTIDE AND PEPTIDOMIMETIC DRUGS

G. L. Amidon and H. J. Lee[1]
College of Pharmacy, The University of Michigan, Ann Arbor, Michigan 48109-1065

KEY WORDS: oral peptide drug absorption, mucosal cell permeability, GI metabolism, liver metabolism, prodrugs, alternative routes

INTRODUCTION

Peptide drug delivery has been of considerable interest for the past 10 to 15 years, due in part to the availability of therapeutic proteins on a commercial scale. The advent of biotechnology and advances in peptide synthesis have made the synthesis of small peptides almost routine. These two factors augmented by the development of receptor-based screening procedures have led to a large number of drug discovery programs focused on peptide-type drugs. For the purposes of this review, we define a peptide-type drug as a drug composed of amino acids or amino acid analogs whose synthesis is based on some analogy with natural peptides or proteins. The focus of this report is principally oral delivery because this is the preferred route of drug administration, though comparative results for other routes are presented.

Oral drug absorption is usually considered to be drug absorbed into the systemic circulation. This is the most relevant definition for the majority of drugs that are active by parenteral administration. However, the most basic definition of drug absorption would be drug absorbed into the gastrointestinal tissue, because once past the intestinal mucosal cell brush border, the drug may be considered to be in the body. This view emphasizes that the processes limiting systemic availability must also

[1]Also affiliated with: TSRL, Inc., 540 Avis Drive, Ann Arbor, Michigan 48108.

0362-1642/94/0415-0321$05.00